

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 18, 2004, 21:02:52 ; Search time 49 Seconds

(without alignments)

8607.685 Million cell updates/sec

Title: US-10-054-536-2

Perfect score: 135

Sequence: 1 atgtccctgtttccatcact.....tctgtgagttccctatctga 747

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 2327084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/uspro.spool.p/US10054536/runat.18062004.173457.13682/app_query.fasta_1.903
-DB=Published Applications AA -QMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10054536 @CGN 1.1.13 @runat.18062004.173457.13682
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LOGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	1321	96.8	248	14	US-10-054-536-28	Sequence 28, Appl
2	1321	96.8	248	15	US-10-388-322-1	Sequence 1, Appl
3	1321	96.8	248	16	US-10-429-160-58	Sequence 58, Appl
4	1228	90.0	228	12	US-09-971-475-1	Sequence 1, Appl
5	454	33.3	375	12	US-10-312-829-9	Sequence 9, Appl
6	442.5	32.4	351	13	US-10-007-408-1	Sequence 1, Appl
7	441	32.3	374	12	US-10-312-829-7	Sequence 7, Appl
8	404	29.6	248	12	US-10-312-829-4	Sequence 4, Appl
9	398.5	28.2	248	12	US-10-312-829-12	Sequence 12, Appl
10	396.5	28.0	243	12	US-10-312-829-96	Sequence 96, Appl
11	394.5	28.9	248	12	US-10-336-603A-100	Sequence 100, App
12	389.5	28.5	259	9	US-09-925-302-473	Sequence 473, App
13	389.5	28.5	259	12	US-09-925-302-473	Sequence 473, App
14	314.5	23.0	153	12	US-10-312-829-10	Sequence 10, Appl
15	304.5	22.3	271	9	US-09-989-723-357	Sequence 357, App
16	304.5	22.3	271	9	US-09-989-723-357	Sequence 357, App
17	304.5	22.3	271	9	US-09-989-723-357	Sequence 357, App
18	304.5	22.3	271	9	US-09-989-723-357	Sequence 357, App
19	304.5	22.3	271	9	US-09-989-731-357	Sequence 357, App
20	304.5	22.3	271	9	US-09-989-732-357	Sequence 357, App
21	304.5	22.3	271	9	US-09-991-073-357	Sequence 357, App
22	304.5	22.3	271	9	US-09-990-442-357	Sequence 357, App
23	304.5	22.3	271	9	US-09-991-163-357	Sequence 357, App
24	304.5	22.3	271	9	US-09-993-604-357	Sequence 357, App
25	304.5	22.3	271	9	US-09-990-456-357	Sequence 357, App
26	304.5	22.3	271	9	US-09-989-721-357	Sequence 357, App
27	304.5	22.3	271	9	US-09-992-598-357	Sequence 357, App
28	304.5	22.3	271	9	US-09-989-293A-357	Sequence 357, App
29	304.5	22.3	271	9	US-09-989-735-357	Sequence 357, App
30	304.5	22.3	271	9	US-09-990-444-357	Sequence 357, App
31	304.5	22.3	271	9	US-09-991-181-357	Sequence 357, App
32	304.5	22.3	271	9	US-09-989-730-357	Sequence 357, App
33	304.5	22.3	271	9	US-09-990-436-357	Sequence 357, App
34	304.5	22.3	271	9	US-09-993-687-357	Sequence 357, App
35	304.5	22.3	271	10	US-09-989-734-357	Sequence 357, App
36	304.5	22.3	271	10	US-09-997-653-357	Sequence 357, App
37	304.5	22.3	271	10	US-09-993-667-357	Sequence 357, App
38	304.5	22.3	271	10	US-09-997-428-357	Sequence 357, App
39	304.5	22.3	271	10	US-09-997-666-357	Sequence 357, App
40	304.5	22.3	271	10	US-09-990-438-357	Sequence 357, App
41	304.5	22.3	271	10	US-09-990-562-357	Sequence 357, App
42	304.5	22.3	271	10	US-09-990-711-357	Sequence 357, App
43	304.5	22.3	271	10	US-09-989-726-357	Sequence 357, App
44	304.5	22.3	271	10	US-09-998-156-357	Sequence 357, App
45	304.5	22.3	271	10	US-09-990-437-357	Sequence 357, App

ALIGNMENTS

RESULT 1

US-10-054-536-28
; Sequence 28, Application US/10054536
; Publication No. US20030162248A1

; GENERAL INFORMATION:

; APPLICANT: Wakamiya, No. US20030162248A1utaka

; TITLE OF INVENTION: RECOMBINANT HUMAN MANNAN-BINDING PROTEINS AND PROCESS

; FILE OF INVENTION: FOR PRODUCING THE SAME

; FILE REFERENCE: 19036/36614

; CURRENT APPLICATION NUMBER: US/10/054,536

; CURRENT FILING DATE: 2002-01-22

; PRIOR APPLICATION NUMBER: PCT/JP98/03311

; PRIOR FILING DATE: 1998-07-23

; PRIOR APPLICATION NUMBER: JP 10-11864

; PRIOR FILING DATE: 1998-01-23

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 28

; LENGTH: 248

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-054-536-28

Alignment Scores:

APPLICANT: Ward, Teresa R
APPLICANT: Mao, Mao
APPLICANT: Linsley, Peter S
APPLICANT: Lund, Lund
TITLE OF INVENTION: LXR Ligand Induced Genes and Proteins
FILE REFERENCE: RS0200
CURRENT APPLICATION NUMBER: US/10/429,160
CURRENT FILING DATE: 2003-05-02
PRIOR APPLICATION NUMBER: US 60/377,714
PRIOR FILING DATE: 2002-05-03
NUMBER OF SEQ ID NOS: 111
SOFTWARE: PatentIn version 3.1
SEQ ID NO 58
LENGTH: 248
TYPE: PRT
ORGANISM: Homo sapiens
US-10-429-160-58

Alignment Scores:
Pred. No.: 1,03e-110 Length: 248
Score: 1321.00 Matches: 248
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.78% Indels: 0
DB: 16 Gaps: 0

US-10-054-536-2 (1-747) x US-10-429-160-58 (1-248)

QY 1 ATGTCCTGTTTCCATCCTCCTCTCTCTGAGTATGTCGAGCGTCTTACTCA 60
DB 1 MetSerLeuPheProSerLeuProLeuLeuLeuSerMetValAlaAlaSerTyrSer 20
QY 61 GAAACTGTGACCTGTGAGGATGCCAAAGACTGCGCTGACGTATGCTGTAGCTCT 120
DB 21 GluThrValThrCysGluAspAlaGlnLysThrCysProAlaValIleAlaCysSerSer 40
QY 121 CAGGATCAACGGCTCCACAGCAAGATGGCGGTGATGCGCACCCAGGAGAGAAAGGG 180
DB 41 ProGlyIleAsnGlyPheProGlyLysAspGlyArgAspGlyThrLysGlyGluLysGly 60
QY 181 GAACAGGCCAAGGGCTCAGAGGCTTACAGGGCCCCCTCGAAAGTTGGGGCTTCCAGGA 240
DB 61 GluProGlyGlnGlyLeuArgGlyLeuGlnGlyProProGlyLysLeuGlyProProGly 80
QY 241 AATCCAGGGCTTCTGGGTCAACAGGACCAAGGCCCAAGAGGAGACCTCGAAAGT 300
DB 81 AsnProGlyProSerGlySerProGlyProLysGlyGlnLysGlyAspProGlyLysSer 100
QY 301 CCGATGCTATAGTAGCTGGCTGCTCAGAGAGAAAGCTCTGCAACAGAGAAATGCGCA 360
DB 101 ProAspGlyAspSerSerLeuAlaAlaSerGluArgLysAlaLeuGlnThrGluMetAla 120
QY 361 CGTATCAAAAGTGGCTGACCTCTCTCTGGGCAACAAAGTTGGGAAACAAAGTTCTTCTG 420
DB 121 ArgIleLysLysTrpLeuThrPheSerLeuGlyLysGlnValGlyAsnLysPhePheLeu 140
QY 421 ACCAATCGTGAATATGACCTTTGAAAGAGTGAAGGCCCTGTGTCAAGTTCAGGCC 480
DB 141 ThrAsnGlyGluIleMetThrPheGluLysValLysAlaLeuCysValLysPheGlnAla 160
QY 481 TCTGTGGCCACCCAGGAATGCTGCAGAGAAATGAGCCATTTCAGAAATCTCATCAAGGAG 540
DB 161 SerValAlaThrProArgAsnAlaAlaGluAsnGlyAlaIleGlnAsnLeuIleLysGlu 180
QY 541 GAAGCTTCTCTGGGATCACTGATGAGAGACAGAAAGGGCAGTGTGTGGATCTTGACAGGA 600
DB 181 GluAlaPheLeuGlyIleThrAspGlnLysThrGluGlyGlnPheValAspLeuThrGly 200
QY 601 AATGACTGACCTACACAACTGGAAACGAGGTTGAACCCAAATGCTGGTTCGATGAA 660
DB 201 AsnArgLeuThrTyrThrAsnTrpAsnGluGlyGluProAsnAsnAlaGlySerAspGlu 220
QY 661 GATTGTGTTGCTACTGAAATAATGGCCAGTGAATGAGTCCCCCTGCTCCACCTCCCAT 720

DB 221 AspCysValLeuLeuLysAsnGlyGlnTrpAsnAspValProCysSerThrSerHis 240
QY 721 CTGGCCGTCTGTGAGTTCCTATC 744
DB 241 LeuAlaValCysGluPheProIle 248
RESULT 4
US-09-971-475-1
Sequence 1, Application US/09971475
Publication NO. US20020086817A1
GENERAL INFORMATION:
APPLICANT: Kawasaki, Toshisuke
TITLE OF INVENTION: Anticancer agent
FILE REFERENCE: ADT 308
CURRENT APPLICATION NUMBER: US/09/971,475
CURRENT FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: US 09/468,705
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: PCT/JP98/03697
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: JP 239113/97
PRIOR FILING DATE: 1997-08-21
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens
US-09-971-475-1

Alignment Scores:
Pred. No.: 2,58e-102 Length: 228
Score: 1228.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.96% Indels: 0
DB: 12 Gaps: 0

US-10-054-536-2 (1-747) x US-09-971-475-1 (1-228)

QY 61 GAAACTGTGACCTGTGAGGATGCCAAAGACTGCGCTGACGTATGCTGTAGCTCT 120
DB 1 GluThrValThrCysGluAspAlaGlnLysThrCysProAlaValIleAlaCysSerSer 20
QY 121 CAGGATCAACGGCTCCACAGCAAGATGGCGGTGATGCGCACCCAGGAGAGAAAGGG 180
DB 21 ProGlyIleAsnGlyPheProGlyLysAspGlyArgAspGlyThrLysGlyGluLysGly 40
QY 181 GAACAGGCCAAGGGCTCAGAGGCTTACAGGGCCCCCTCGAAAGTTGGGGCTTCCAGGA 240
DB 41 GluProGlyGlnGlyLeuArgGlyLeuGlnGlyProProGlyLysLeuGlyProProGly 60
QY 241 AATCCAGGGCTTCTGGGTCAACAGGACCAAGGCCCAAGAGGAGACCTCGAAAGT 300
DB 61 AsnProGlyProSerGlySerProGlyProLysGlyGlnLysGlyAspProGlyLysSer 80
QY 301 CCGATGCTATAGTAGCTGGCTGCTCAGAGAGAAAGCTCTGCAACAGAGAAATGCGCA 360
DB 81 ProAspGlyAspSerSerLeuAlaAlaSerGluArgLysAlaLeuGlnThrGluMetAla 100
QY 361 CGTATCAAAAGTGGCTGACCTCTCTCTGGGCAACAAAGTTGGGAAACAAAGTTCTTCTG 420
DB 101 ArgIleLysLysTrpLeuThrPheSerLeuGlyLysGlnValGlyAsnLysPhePheLeu 120
QY 421 ACCAATCGTGAATATGACCTTTGAAAGAGTGAAGGCCCTGTGTCAAGTTCAGGCC 480
DB 121 ThrAsnGlyGluIleMetThrPheGluLysValLysAlaLeuCysValLysPheGlnAla 140
QY 481 TCTGTGGCCACCCAGGAATGCTGCAGAGAAATGAGCCATTTCAGAAATCTCATCAAGGAG 540
DB 141 SerValAlaThrProArgAsnAlaAlaGluAsnGlyAlaIleGlnAsnLeuIleLysGlu 160

```
QY 541 GAAGCCTTCTGGGATCATCTGATGAGAACAGAGAGGCGAGTTTGTGGATCTGACAGGA 600
Db 161 GluAlaPheLeuGlyIleThrAspGluLysThrGluGlnPheValAspLeuThrGly 180
QY 601 AATAGACTCACCTACAACTGGAACGAGGTGAACCCAACTGCTGGTCTGTATGAA 660
Db 181 AsnArgLeuThrTyrThrAsnTrpAsnGluGlyGluProAsnAsnAlaGlySerAspGlu 200
QY 661 GATTGTGATTCTACTGAAAAATGGCCAGTGGAAATGAGTGCCTCCCTGCTCCACCTCCCAT 720
Db 201 AspCysValLeuLeuLeuLysAsnGlyGlnTrpAsnAspValProCysSerThrSerHis 220
QY 721 CTGGCGCTGTGAGTTCCTATC 744
Db 221 LeuAlaValCysGluPheProIle 228

RESULT 5
US-10-312-829-9
; Sequence 9, Application US/10312829
; Publication No. US20040037781A1
; GENERAL INFORMATION:
; APPLICANT: McCormack, Francis X
; TITLE OF INVENTION: Peptides with Antioxidant and Antimicrobial Properties
; FILE REFERENCE: 10738-31
; CURRENT APPLICATION NUMBER: US/10/312,829
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: PCT/US01/21226
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,313
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-312-829-9

Alignment Scores:
Pred. No.: 2,83e-32 Length: 375
Score: 454.00 Matches: 375
Percent Similarity: 55.07% Conservative: 26
Best Local Similarity: 43.61% Mismatches: 82
Query Match: 33.26% Indels: 20
DB: 12 Gaps: 5

US-10-054-536-2 (1-747) x US-10-312-829-9 (1-375)
QY 118 TCTCCAGCATCAACGGCTTCCAGCAAGATGGCGTGATGGCAACAGGGAGAAAG 177
Db 149 AlaProGlyMetGlnGlySerAlaGlyAlaArgGlyLeuAlaGlyProLysGlyGluArg 168
QY 178 GGGGAACACAGGCAAA-----GGGCTCAGAGCTTACAGGCGCCCTCGGAAG 225
Db 169 GlyValProGlyGluArgGlyValProGlyAsnThrGlyAlaAlaGlySerAlaGlyAla 188
QY 226 TTGGGGCTCCAGAAATCCAGGCGCTTCTGGTCCAGGACCAAGGGCGCAAAAGGA 285
Db 189 MetGlyProGlnGlySerProGlyAlaArgGlyProGlyLeuLysGlyAspLysGly 208
QY 286 GACCCTGGA--AAAGTCCGGATGTGATAGTACCTG-----GCTGCTCAGAAAGA 336
Db 209 IleProGlyAspLysGlyAlaLysGlyGluSerGlyLeuProAspValAlaSerLeuArg 228
QY 337 AAAGCT-----CTGCAAAACAGAAATGGCACGTATC 366
Db 229 GlnGlnValGluAlaLeuGlnGlyValGlnHisLeuGlnAlaPheSerGlnTyr 248
QY 367 AAAAGTGGGTGACCTTCTCTCTGGGCAAAACAAGTTGGGAACAAGTTCTTCTTCGACCAAT 426
Db 249 LysLysValGluLeuLeuPheProAsnGlyGlnSerValGlyGluLysIlePheLysThrAla 268
QY 427 GGTGAAATATGACCTTTGAAAAAGTCAAGGCCTTGTGTGTCAGTTCAGGCGCTCTGTG 486
```

```
Db 269 GlyPheValLysProPheThrGluAlaGlnLeuLeuCysThrGlnAlaGlyGlyGlnLeu 288
QY 487 GCCACCCAGGAATGCTGCAGAGATGGAGCAATTCAGAAATCTCATC-----AAG 537
Db 289 AlaSerProArgSerAlaAlaGluAsnAlaAlaLeuGlnGlnLeuValValAlaLysAsn 308
QY 538 GAGGAAGCCTTCTCTGGGCATCTGATGAGAGACAGAAAGGCGAGTTTGTGGATCTGACA 597
Db 309 GluAlaAlaPheLeuSerMetThrAspSerLysThrGluGlyLysPheThrTyrProThr 328
QY 598 GGAATATAGACTGACCTACACAACTGGAACGAGGTGAACCCCAACAATCTGTTCTGAT 657
Db 329 GlyGluSerLeuValTyrSerAsnTrpAlaProGlyGluProAsnAspGlyGlySer 348
QY 658 GAAGATTGTATTCTACTGAAAAATGGCCAGTGGAAATGAGTGCCTCCCTGCTCCACCTCC 717
Db 349 GluAspCysValGluIlePheThrAsnGlyLysTrpAsnAspArgAlaCysGlyGluLys 368
QY 718 CATCTGGCGCTGTGAGTTC 738
Db 369 ArgLeuValValCysGluPhe 375

RESULT 6
US-10-007-408-1
; Sequence 1, Application US/10007408
; Publication No. US20020168627A1
; GENERAL INFORMATION:
; APPLICANT: WAKAMIYA, NO. US20020168627A1utaka
; TITLE OF INVENTION: RECOMBINANT CONGLUTININ AND PRODUCING METHOD THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/007,408
; FILING DATE: 08-No. US20020168627A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/02035
; FILING DATE: 02-OCT-1995
; APPLICATION NUMBER: JPA - 209698
; FILING DATE: 17-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 19036/34546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-007-408-1

Alignment Scores:
Pred. No.: 3.05e-31 Length: 351
Score: 442.50 Matches: 91
```



```

; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: 09/746,491
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 10/055,569
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 100
; LENGTH: 248
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-336-603A-100

```

Alignment Scores:		
Pred. No.:	6.18e-27	248
Score:	394.50	96
Percent Similarity:	50.57%	96
Best Local Similarity:	36.78%	30
Query Match:	298.9%	39
DB:	12	11
	1.2	11
	Length:	
	Matches:	
	Conservative:	
	Mismatches:	
	Indels:	
	Gaps:	

US-10-054-536-2 (1-747) x US-10-336-603A-100 (1-248)

22	CTCTCTCTTCTCTCTGAGTATGGTGCACGCTTACTTACTCAGAACTGTGACCTGTGAGGAT	81
Db	: : : : : : : : : :	
5	ProLeuAlaLeuAsnLeuIleLeuMetAlaIaSerGlyAlaValCysGluValLysAsp	24
QY	82 GCCCAAAAGACCTGCCCTCGCATGTATGCCTGTAGCTCTCCAGGCATCAACGGCTTCCCA	141
Db	: : : : : : : : : :	
25	ValCysValGlySerProGlyIle-----ProGlyThrProGlySerHisGlyLeuPro	42
QY	142 GCCAAAGATGGCGTATGGCACCAAGGAGAAAGGGGNAACAGAGCCCAAGGCTCAGA	201
Db	: : : : : : : : : :	
43	GlyArgAspGlyArgAspGlyValLysGlyAspProGlyPro-----	59
QY	202 GCGTTACAGGGCCCCCTCGAAAGTTGGGG---CTCCAGGAAAT-----CCAGGG	249
Db	: : : : : : : : : :	
60	-----MetGlyProProGlyGluMetProCysProProGlyIleAsnAspGlyLeuProGly	77
QY	250 CTTCTGGGTACACGAGCAAAAGGGCCAAAGAGACCCCTGGA---AAAAGTCCGGAT	306
Db	: : : : : : : : : :	
78	AlaProGlyIleProGlyGluCysGlyGluLysGlyGluProGlyGluArgGlyProPro	97
QY	307 GGT-----GATAGTAGCTGCTGCCTCA-----GAAAGAAA	339
Db	: : : : : : : : : :	
98	GlyLeuProAlaHisLeuAspGluGluAlaThrLeuHisAspPheArgHisGln	117
QY	340 GCTCTGCAACAGAAATGGCACGCTATCAAAAAGTGGCTGACCTTCTCTCTGGCCAAACA	399
Db	: : : : : : : : : :	
118	IleLeuGlnThrArgGlyAla-----LeuSerLeuGln	128
QY	400 -----GTGGGAACAAGTTCTTCTGACCAATGGTGAATAATGACCTTT	444
Ddb	: : : : : : : : : :	
129	GlySerIleMetThrValGlyGluLysValPheSerSerAsnGlyGlnSerIleThrPhe	148
QY	445 GAAAAGTGAAGGCTTGTGTCTCAAGTTCAGGCCTCTGTGGCCACCCCGAGGAATGCT	504
Db	: : : : : : : : : : : : : : :	
149	AspAlaIleGlnAlaCysAlaArgAlaGlyArgIleAlaValProArgAsnPro	168
QY	505 GCAGAGAATGAGGCATTCAGAATCTCATCAAGAG-----GAAAGCTTCTGGGC	555
Db	: : : : : : : : : :	
169	GluGluAsnGluAlaIleAlaSerPheValLysLysTyrAsnThrTyrAlaTyrValGly	188
QY	556 ATCACTGATGAGACAGAGGCGAGTTTGTGGATCTGACGAAATAGACTGACCTTAC	615
Db	: : : : : : : : : : : : : : :	
189	LeuThrGluGlyProSerProGlyAspPheArgTyrSerAspGlyThrProValAsnTyr	208
QY	616 ACAAACTGGACAGAGGTGAACCCACAAATGCTGGTTCTGATGAAGATTGTGTATTGCTA	675
Db	: : : : : : : : : : : : : : :	
209	ThrAsnTrpTyrArgGlyGluProAlaGlyArgGlyLys---GluGlnCysValGluMet	227
QY	676 CTGAAAATGGCCAGTGAATGACGTCCCTCCACCTCCCATCTGCCCTCTGTGAG	735
Db	: : : : : : : : : : : : : : :	

```

Db      228  TyrThrAspGlyGlnTrpAsnAspArgAsnCysLeuTyrSerArgLeuThrIleCysGlu 247
Qy      736  TTC 738
      |||
Db      248  Phe 248

RESULT 12
US-09-925-302-473
; Sequence 473, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0

```

```

; LENGTH: 259
; TYPE: PRT

```

ORGANISM: Homo sapiens

; FEATURE:

```

; NAME/KEY: SITE
: LOCATION: (20)

```

; LOCATION: (20)
; OTHER INFORMATION: Xaa

; NAME/KEY: SITE

```

; LOCATION: (61)
; OTHER INFORMATION: Y22

```

; NAME/KEY: SITE

; LOCATION: (234)

OTHER INFORMATION: Xaa
11C-09-82E-202-473

03-09-923-302-413

Alignment Scores:

Pred. No.:

Score: 38
Percent Similarity: 50

Best Local Similarity: 36

Query Match: 28

DE: 5

US-10-054-536-2 (1-747) x 1

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

22 CCTCTCCTTCTC

Db
16 ProLeuAlaLeu*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 82 GCCCAAAGACCT

Db 36 ValCysValGlyS

100

Qy 142 GGCAAAGATGGG

Db
54 GlvArqAspGlyA

100

QY 202 GGCTTACAGGGCC

D^b 71 -----MetGlyP

100

Qy 250 CCTTCTGGGTCACT

Db 89 AlaProGlyValFF

.....

Qy 307 GGT-----
|||

Db 109 GlyLeuProAlaH

3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842

QY 340 GCTCTGCAACACAG

Db	129	IleLeuGlnThrArgGlyAla-----LeuSerLeuGln	139
Qy	400	-----GTTGGGAACAAGTCTTCTCGACCAATGGTGAATAATGACCTTT	444
Db	140	GlySerIleMetThrValGlyGluLysValPheSerSerAsnGlyGlnSerIleThrPhe	159
Qy	445	GA AAAAGTGAAGGCCTTGTTGTCGAAGTCCAGGCCTCTGTGGCCACCCACAGGAATGCT	504
Db	160	AspAlaIleGlnGluAlaCysAlaAargAlaGlyArgIleAlaValProArgAsnPro	179
Qy	505	GCAGGAATGGAGCCATTCAGAAATCTCATCAAGGAG-----GAAGCCTTCCTGGGC	555
Db	180	GluGluAsnGluAlaIleAlaSerPheValLysLysTyrAsnThrTyrAlaTyrValGly	199
Qy	556	ATCACTGATGAGACAGACAGAGGGCAGTTGTGGATCTGACAGGAAATAGACTGACCTAC	615
Db	200	LeuThrGluGlyProSerProGlyAspPheArgTyrSerAspGlyThrProValAsnTyr	219
Qy	616	ACAACTCGAACGAGGGTGACCCCAACAATGCTGGTCTTGATGAAGATGTGTATTGCTA	675
Db	220	ThrAsnTrpTyrArgGlyGluProAlaGlyArgGlyLys---Glu***CysValGluMet	238
Qy	676	CTGAAAAATGGCCAGTGGAAATGACGTCCCTGCTCCACCTCCCATGTCGCCGCTGTGTAG	735
Db	239	TyrThrAspGlyGlnTrpAsnAspArgAsnCysLeuTyrSerArgLeuThrIleCysGlu	258
Qy	736	TTC	738
Db	259	Phe	259

RESULT 13

```

US-09-925-302-473
; Sequence 473, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 473
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (20)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (61)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (234)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-473

Alignment Scores:
Pred. No.: 1,77e-26 Length: 259
Score: 389.50 Matches: 95
Percent Similarity: 50.19% Conservative: 36
Best Local Similarity: 36.40% Mismatches: 91
Query Match: 28.53% Indels: 39
DB: 12 Gaps: 11

US-10-054-536-2 (1-747) x US-09-925-302-473 (1-259)

```

```

; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-312-829-10

Alignment Scores:
Pred. No.:          9,48e-20          Length:          153
Score:             314.50           Matches:          62
Percent Similarity: 59.26%           Conservative:    18
Best Local Similarity: 45.93%          Mismatches:       52
Query Match:       23.04%             Indels:           3
DB:                12                Gaps:             1

US-10-054-536-2 (1-747) x US-10-312-829-10 (1-153)
QY 343 CTGCAACAGAAATGCGACGTATCAAAAAGTCGTGACCTCTCTGGGCAACAAGTT 402
Db 19 LeuGlnAlaAlaPheSerGlnTyrlsValGluLeuPheProAsnGlyGlnSerVal 38
QY 403 GGAACAAATGTTCTCTGACCAATGGTGAATAATGACCTTTGAAAAAGTGAAGCCCTTG 462
Db 39 GlyGluLysIlePheIleThrAlaGlyPheValLysProPheThrGluAlaGlnLeuLeu 58
QY 463 TGTGTCAGATCCAGGCTCTGTGGCCACCCAGCAATGTCGACAGAAATGGAGCCATT 522
Db 59 CysThrGlnAlaGlyGlyGlnLeuAlaSerProArgSerAlaAlaGluAsnAlaAlaLeu 78
QY 523 CAGAAATCTCATC-----AAGGAGGAAGCTTCTGGGCATCACTGACGAGAGACA 573
Db 79 GlnGlnLeuValValAlaLysAsnGluAlaAlaPheLeuSerMetThrAspSerLysThr 98
QY 574 GAAGGCGATGTTGTGGATCTGACAGAAATAGACTGACCTACACAACTGGAACGAGGTT 633
Db 99 GluGlyLysPheThrTy-ProThrGlySerLeuValTySerAsnTrpAlaProgly 118
QY 634 GAACCCAAATGCTGGTTCGATGACAAATGTGTATGCTACTGAAAAATGGCCAGTGG 693
Db 119 GluProAsnAspGlyGlySerGluAspCysValGluIlePheThrAsnGlyLysTrp 138
QY 694 AATGACGTCCTCTGCTCCACCTCCATCTGGCGCTCTGTGAGTTC 738
Db 139 AsnAspArgAlaCysGlyGluLysArgLeuValValCysGluPhe 153

RESULT 15
US-09-989-722-357
; Sequence 357, Application US/09989722
; Patent No. US20020072067A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gotttsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC63
; CURRENT APPLICATION NUMBER: US/09/989,722
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
```

;
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25

;
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 8,67e-19 Length: 271
Score: 304.50 Matches: 84
Percent Similarity: 44.73% Conservative: 39
Best Local Similarity: 30.55% Mismatches: 96
Query Match: 22.31% Indels: 56
DB: 9 Gaps: 9

US-10-054-536-2 (1-747) x US-09-989-722-357 (1-271)

QY 9 GTTTCATCACTCCCTCTCTCTGAGTATGGTGGCAGCGCTCTTACTCAGAAACTGT 68
Db 17 LeuSerLeuLeuProSerGlyHisProGlnProAlaGly----- 29
QY 69 GACCTGTGAGGATGCCCAAAAGACCTGCCCTGCAGTGATTGCTGTAGCTCTCCAGGCAT 128
Db 30 -----AspAspAlaCysSer-ValGlnIleLeuValProGlyLe 42
QY 129 C-----AACGGCTTCCAGGCAAAAGATGGGGCTGATGG 161
Db 42 uLysGlyAspAlaGlyGluLysGlyAspLysGlyAlaProGlyArgProGlyArgValGl 62
QY 162 CACCAAGGGAGAAAAGGGGGAACACAGGC---CAAGGGCTCAGAGGCTTACAGGGCCCCC 218
Db 62 YProThrGlyGluLysGlyAspMetGlyAspLysGlyGlnLysGlySerValGlyArgHi 82
QY 219 TGGAAAAGTTGGGGCCT-----CCAGGAATCCAGGSCCTTCTGGGTACCAGGACC 269
Db 82 sGlyLysIleGlyProIleGlySerLysGlyGluLysGlyAspSerGlyAspIleGlyPr 102
QY 270 AAAGGGCCAAAAGGAGACCTCGA----- 294
Db 102 oProGlyProAsnGlyGluProGlyLeuProCysGlyCysSerGlnLeuArgLysAlaIl 122
QY 295 -AAAAGTCCGGATGGTGATAGTAGCTGGCTGCCTCAGAAAGAAAAGCTCTGCAAAACAGA 353
Db 122 eGlyGluMetAspAsnGlnValSerGlnLeuThrSerGluLeuLysPheIleLysAsnAl 142
QY 354 AATGGCACGTATCAAAAAGTGGGTGCTCTCTCTGGGCAAAACAGTTGGGACACAGTT 413
Db 142 aValAlaGlyVal-----ArgGluThrGluSerLysIl 153
QY 414 CTTCTGACCAATGGTGAATAATGACCTTTGAAAAGTGAAGCCCTTGTGTGTCAGATT 473
Db 153 eTyrLeuLeuValLysGluLysArgTyrAlaAspAlaGlnLeuSerCysGlnGlyAr 173

474	QY	CRAGGCCTCTGTGGCCACCCCCAGGAATGTCGAGAGAATCGAGCCATTCAGAATCTCAT	533
173	Db	gGlyGlyThrLeuSerMetProLysAspGluLaalaaAsnGlyLeuMetAlaLaTyLe	193
534	QY	CAAGGAG-----GAAGCTTCTCGGGCATCACTCATGAGAAGACAGAGGGCA	581
193	Db	uAlaGlnAlaGlyLeuAlaargValPheIleGlyLeaAsnAspLeuGluLysGluGlyAl	213
582	QY	GTTTGTG-----GATCTGCAGGAATAGACTGACCTACACAACTGGACGAGGGTGA	635
213	Db	aPheValTySerAspHisSerProMetArg---ThrPheAsnLysTrpArgSerGlyGly	232
636	QY	ACCCACAATGCTGGTTCTCATGAAGATTGTGTATTGCTACTGAAAAATGCCAGTCGAA	695
232	Db	uProAsnAlaTyAspGluGluAspCysValGluMetValAlaSerGlyGlyTrpAs	252
696	QY	TGACGTCCCTGCTCCACCTCCCATCTGGCGCTGTGAGTTC	738
252	Db	naspValAlaCysHisThrThrMetTyPheMetCysGluPhe	266

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 18, 2004, 20:57:11 ; Search time 21 Seconds
(without alignments)
6843.340 Million cell updates/sec

Title: US-10-054-536-2
Perfect score: 1365
Sequence: 1 atgtccctgttccatcact.....tctgtgagttccctatctga 747

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10054536/runat.18062004.173456.13600/app.query.fasta.1.903
-DB=PIR 78 -QPMF=fastan -SUFFIX=tdr -MINMATCH=0.1 -LOOPEXT=0 -TRANS=human40.cdi -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10054536 @CGN 1.1.38 @runat.18062004.173456.13600 -NCPV=6 -ICPV=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBPLOCK=100 -LONGIOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 78 : *
1: pir1 : *
2: pir2 : *
3: pir3 : *
4: pir4 : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1321	96.8	248	1 LNHUMC	mannose-binding le
2	790	57.9	244	1 LNMWSC	mannose-binding le
3	785	57.5	244	1 LNMTC	mannose-binding le
4	711.5	52.1	239	1 LNMWA	mannose-binding le
5	705.5	51.7	238	1 LNMWA	mannose-binding le
6	454	33.3	375	1 A45225	pulmonary surfacta
7	441.5	32.3	371	1 JN0450	conglutinin precu
8	441	32.3	374	1 A42046	surfactant protein
9	440.5	32.3	371	2 I45878	conglutinin - bovi
10	418	30.6	301	2 A53570	collectin-43 - bov
11	417.5	30.6	247	1 LNRBPS	pulmonary surfacta
12	416.5	30.5	369	2 S33603	surfactant protein
13	404	29.6	248	1 LNRTPS	pulmonary surfacta
14	397.5	29.1	248	1 LNHUPI	pulmonary surfacta

15	393.5	28.8	248	1 LNHUP6	pulmonary surfacta
16	393	28.8	248	2 A48853	pulmonary surfacta
17	385.5	28.2	248	1 LNHUPS	pulmonary surfacta
18	385.5	28.2	248	2 I51921	pulmonary surfacta
19	368.5	27.0	248	1 LNDGFS	scavenger receptor
20	278.5	20.4	742	2 JC7595	collectin liver 1
21	272.5	20.0	277	2 JC7903	IgE Fc receptor, l
22	201.5	14.8	331	1 LNMSE	hepatic lectin - c
23	199.5	14.6	207	1 LNCHL	IgE Fc receptor II
24	196.5	14.4	309	1 S34198	complement C1q B c
25	194.5	14.2	253	2 I49560	complement protein
26	192.5	14.1	253	2 S49158	gelatin-binding 28
27	192	14.1	244	2 JC4708	collagen alpha 3(I
28	189	13.8	245	2 S19018	collagen alpha 3(I
29	188	13.8	1670	1 CGHU3B	ficollin-A precursor
30	187	13.7	334	2 JC5980	collagen alpha 1(X
31	187	13.7	1747	2 A45974	collagen alpha 2(I
32	187	13.7	1763	2 S16366	collagen alpha 1(X
33	187	13.7	1857	2 S31212	collagen alpha 1(X
34	187	13.7	1888	2 S78476	complement subcomp
35	186.5	13.7	245	1 C1HUQC	collagen alpha 4(I
36	186	13.6	1690	1 CGHULB	tetranectin precu
37	185.5	13.6	202	1 TTHUN	collagen alpha 1(X
38	185.5	13.6	330	2 S46657	brevican precursor
39	185	13.6	912	2 A54423	complement subcomp
40	184	13.5	253	1 C1HUQB	complement subcomp
41	183.5	13.4	245	1 C1HUQA	collagen alpha 3(I
42	183	13.4	675	2 S20819	collagen alpha 2(I
43	183	13.4	688	2 A53330	complement subcomp
44	182.5	13.4	246	2 S29328	collagen alpha 1(I
45	182.5	13.4	673	1 CGB06C	collagen alpha 1(I

ALIGNMENTS

RESULT 1

LNHUMC

N;Alternate names: mannan-binding protein
mannose-binding lectin precursor [validated] - human

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1989 #sequence revision 30-Sep-1991 #text change 08-Dec-2000

C;Accession: JLO115; S05641; A34978; JLO027; JX0319; PC2188; A32266

R;Sastry, K.; Herman, G.A.; Day, L.; Deignan, E.; Bruns, G.; Morton, C.C.; Ezekowitz, R. J. Exp. Med. 170, 1175-1189, 1989

A;Title: The human mannose-binding protein gene. Exon structure reveals its evolutionary

A;Reference number: JLO115; MUID:90010778; PMID:2477486

A;Accession: JLO115

A;Molecule type: DNA

A;Residues: 1-248 <SAS>

A;Cross-references: EMBL:X15422; NID:g34486; PIDN:CAA33462.1; PID:g34487

R;Taylor, M.E.; Brickell, P.M.; Craig, R.K.; Summerfield, J.A.

Biochem. J. 262, 763-771, 1989

A;Title: Structure and evolutionary origin of the gene encoding a human serum mannose-b-

A;Reference number: S05641; MUID:90073571; PMID:2590164

A;Accession: S05641

A;Molecule type: DNA

A;Residues: 1-248 <TAY>

A;Cross-references: EMBL:X15954; NID:g34480; PIDN:CAA34079.1; PID:gl1212951

A;Accession: A34978

A;Molecule type: protein

A;Residues: 'X',22-24,'X',26,'X',28-31,'X',33-34,'X',36,'XXXX',41-50 <TAY2>

R;Ezekowitz, R.A.B.; Day, L.E.; Herman, G.A.

J. Exp. Med. 167, 1034-1046, 1988

A;Title: A human mannose-binding protein is an acute-phase reactant that shares sequence

A;Reference number: JLO027; MUID:86171281; PMID:2450948

A;Accession: JLO027

A;Molecule type: mRNA

A;Residues: 1-2,'C',4,'T',8,'S',10-57,'R',59-60,'GT',63-106,'PGCLRK',113,'SSANRNCTYQ',

R;Kurata, H.; Samnol, T.; Kozutsumi, Y.; Yokota, Y.; Kawasaki, T.

J. Biochem. 115, 1148-1154, 1994

A;Title: Structure and function of mannan-binding proteins isolated from human liver an

A;Reference number: JX0319; MUID:95073978; PMID:7982896

A;Accession: JX0319

JS-10-054-536-2 (1-747) x LNMSMC (1-244)

Db 137 SerSerValArgMetProLeuAenArgAlaLyLeuAenCysSerGluLeuGlnGly 156
 QY 481 TCTGTGGCCACCCAGCAATGTCAGAGAAATGGAGCCATTGAGAAATCTCATCAGGAG 540
 Db 157 ThrValAlaThrProArgAsnAlaGluGluAenArgAlaGlnAsnValAlaLysAsp 176
 QY 541 GAAGCTTCTGGGATCACTGATGAGAGACAGAGGCGAGTTTGTGATCTGCACAGA 600
 Db 177 ValAlaPheLeuGlyLeuThrAspGlnArgThrGluAenValPheGluAspLeuThrGly 196
 QY 601 AATAGACCTGACCTACACAACTGGACAGGCTGAACCAACAATGCTGCTGATGAA 660
 Db 197 AsnArgValArgTyrThrAsnTrpAsnGluGlyGluProAsnAsnValGlySerGlyGlu 216
 QY 661 GATTGTGATTCTACTGAAAATGCGCAGTGGATGAGTCCCTGCTGCTCCACCTCCCAT 720
 Db 217 AsnCysValValLeuLeuThrAsnGlyTyrTrpAsnAspValProCysSerAspSerPhe 236
 QY 721 CTGGCCGCTGTGAGTTC 738
 Db 237 LeuValValCysGluPhe 242

RESULT 4

LNM5MA

mannose-binding lectin A precursor - mouse
 N:Alternate names: Ra-reactive factor P28b; serum mannan-binding protein
 C:Species: Mus musculus (house mouse)
 C>Date: 18-Jun-1993 #sequence revision 02-Jun-1994 #text_change 16-Jun-2000
 C:Accession: A46466; B42574; I48650
 R:Sastry, K.; Zahedi, K.; Lelias, J.M.; Whitehead, A.S.; Ezekowitz, R.A.
 J. Immunol. 147, 692-697, 1991

A>Title: Molecular characterization of the mouse mannose-binding proteins. The mannose-b
 A:Reference number: A46466; MUID:91302823; PMID:1712818
 A:Accession: A46466

A:Molecule type: mRNA

A:Residues: 1-239 <SAS1>

A:Cross-references: GB:S42292; NID:91679939; PIDN:AAB19342.1; PID:g233016

A:Experimental source: Inbred CBA/J, acute phase liver library, PTZ 19 vector

A>Note: sequence extracted from NCBI backbone (NCBIN:42292, NCBI:P42293)

R:Kuge, S.; Inara, S.; Watanabe, E.; Watanabe, M.; Takishima, K.; Suga, T.; Mamiya, G.;
 Biochemistry 31, 6943-6950, 1992

A>Title: cDNAs and deduced amino acid sequences of subunits in the binding component of

A:Reference number: A42574; MUID:92345256; PMID:1637828

A:Accession: B42574

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA; protein

A:Residues: 1-239 <KUG>

A:Cross-references: GB:D11441; NID:g220587; PIDN:BAA02006.1; PID:g220588

A:Experimental source: BALB/c, liver

A>Note: sequence extracted from NCBI backbone (NCBI:P110142)

R:Sastry, R.; Wang, J.S.; Brown, D.C.; Ezekowitz, R.A.; Tauber, A.I.; Sastry, K.N.
 Mamm. Genome 6, 103-110, 1995

A>Title: Characterization of murine mannose-binding protein genes Mb11 and Mb12 reveals

A:Reference number: I48650; MUID:95284466; PMID:7766991

A:Accession: I48650

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-239 <SAS2>

A:Cross-references: EMBL:U09010; NID:g773278; PIDN:AAA82009.1; PID:g773280

C:Comment: Mannose-binding lectins are opsonins that are important in host defense again

C:Genetics:

A:Gene: Mb11

A:Introns: 52/1; 91/1; 116/1

C:Superfamily: mannose-binding lectin; C-type lectin homology

C:Keywords: acute phase; calcium; hydroxyproline; lectin; liver; plasma

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-238/product: mannose-binding lectin A #status predicted <MAT>

F:36-88/Region: collagen-like

F:119-235/Domain: C-type lectin homology <LCH>

Alignment Scores:

Pred. No.: 1,88e-55 Length: 239

Score: 711.50 Matches: 144

Percent Similarity: 69.48% Conservative: 29
 Best Local Similarity: 57.83% Mismatches: 63
 Query Match: 52.12% Indels: 13
 DB: 1 Gaps: 4

US-10-054-536-2 (1-747) x LNM5MA (1-239)

QY 1 ATGTCCCTGTTTCCATCACTCCCTCTCTTCCTGAGTATGGTGACACGCTTACTCA 60
 Db 1 MetLeuLeuProLeuLeuPro---ValLeuLeuCysValValSerValSerSer 19
 QY 61 GAAACTGTGACCTGTGAGGATGCCCAAGACCTGCCCTGCAGTGTATCCCTGTACTCT 120
 Db 20 GlySerGlnThrCysGluAspThrLeuIysThrCys---SerValIleAlaCys----- 36
 QY 121 CCAGGCATCAACGGCTTCCCAAGCAAGATGGCGCTGATGGCAAGAGGAGGAGGAGGG 180
 Db 37 -----GlyArgAspGlyArgAspGlyProIysGlyGlyGlyGly 49
 QY 181 GAACCCAGCCAGGCTCAGAGGCTTACAGGCCCCCTCGAAAGTTGGGGCTCCAGGA 240
 Db 50 GluProGlyGlnGlyLeuArgGlyLeuGlnGlyProGlyLysLeuGlyProGly 69
 QY 241 AATCCAGGCGCTTCTGGGCTCACCAGGACCAAGGCGCCAAAGGAGGAGCCCTGGAAAAAGT 300
 Db 70 SerValGlySerProGlySerProGlyProIysGlyGlnLysGlyAspHisGlyAspAsn 89
 QY 301 CCGAGTGTGTATGATGAGCTGGCT-----GCCTCAGAAAGAAAGAAAGCTCTGCAACAGAA 354
 Db 90 ArgAlaIleGluGluLysLeuAlaAsnMetGluAlaGluIleArgIleLeuLysSerLys 109
 QY 355 ATGGCAGCTATCAAAAAGTGGCTGACCTTCTCTGGGCAACAAAGTTGGGAACAAAGTTC 414
 Db 110 LeuGlnLeuThrAsnLysLeuHisAlaPheSerMetGlyLysLysSerGlyLysLysLeu 129
 QY 415 TTCCTGACCAATGTGAAATAATGACCTTTCAAAAGTGAAGGCTCTGTGTCAAGTTC 474
 Db 130 PheValThrAsnHisGluLysMetProPheSerLysValLysSerLysCysThrGluLeu 149
 QY 475 CAGGCTCTGTGGCCACCCCGAGAAATGCTCAGAGAAATGAGGCAATTCAGAATCTCATC 534
 Db 150 GlnGlyThrValAlaIleProArgAsnAlaGluGluAsnLysAlaIleGlnGluValAla 169
 QY 535 AAGGAGGAAGCTTCTGGGCTCAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594
 Db 170 ThrGlyIleAlaPheLeuGlyIleThrAspGluAlaThrGluGlyGlnPheMetTyrVal 189
 QY 595 ACAGGAATAGACTGACCTACACAACTGGAAACGAGGCTGAACCAACAAATGCTGCTTCT 654
 Db 190 ThrGlyGlyArgLeuThrTyrSerAsnTrpLysLysAspGluProAsnAsnHisGlySer 209
 QY 655 GATGAAAGTGTGTATGCTTACTGAAATAATGGCCAGTGAATGAGTCCCTGCTGCCACC 714
 Db 210 GlyGluAspCysValIleIleLeuAspAsnGlyLeuTrpAsnAspIleSerCysGlnAla 229
 QY 715 TCCCATCTGGCCGCTGTGTGAGTCCCT 741
 Db 230 SerPheLysAlaValCysGluPhePro 238

RESULT 5

LNR7MA

mannose-binding lectin A precursor - rat

N:Alternate names: serum mannan-binding protein

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text_change 22-Jun-1999

C:Accession: B24791; A29556; A27799

R:Drickamer, K.; Dordal, M.S.; Reynolds, L.

J. Biol. Chem. 261, 6878-6887, 1986

A>Title: Mannose-binding proteins isolated from rat liver contain carbohydrate-recognit
 rotein.

A:Reference number: A24791; MUID:86196130; PMID:3009480

A:Accession: B24791

A:Molecule type: mRNA

A;Residues: 1-238 <DRI>
R;Drickamer, K.; McCreary, V.
J. Biol. Chem. 262, 2582-2589, 1987
A;Title: Exon structure of a mannose-binding protein gene reflects its evolutionary relationship
A;Reference number: A29556; MUID:87137502; PMID:3029088
A;Accession: A29556
A;Molecule type: DNA
A;Residues: 1-155, 'K', 157-238 <DR2>
A;Cross-references: GB:M14104; GB:M14105; NID:g205259; PIDN:AAA98781.1; PID:g205261
A;Note: The codon AAG for residue 156 is inconsistent with the authors' statement that Cys is at this position
R;Ikeda, K.; Samoh, T.; Kawasaki, N.; Kawasaki, I.
J. Biol. Chem. 262, 7451-7454, 1987
A;Title: Serum lectin with known structure activates complement through the classical pathway
A;Reference number: A27799; MUID:87222358; PMID:3584121
A;Accession: A27799
A;Molecule type: protein
A;Residues: 18-42 <IKE>
C;Comment: Mannose-binding lectins are opsonins that are important in host defense against bacteria
C;Comment: This plasma protein binds mannose and N-acetylglucosamine and can activate complement
C;Comment: The molecule consists of approximately 20 identical chains linked by disulfide bonds
C;Genetics:
A;Introns: 51/1; 90/1; 115/1
A;Superfamily: mannose-binding lectin; C-type lectin homology
C;Keywords: acute phase; calcium; hydroxylysine; hydroxyproline; lectin; liver; plasma
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-238/Product: mannose-binding lectin A #status experimental <WAT>
F;36-88/Region: collagen-like
F;85-87/Region: cell attachment (R-G-D) motif
F;118-234/Domain: C-type lectin homology <LCH>
F;61,67,73/Modified site: 4-hydroxyproline (Pro) #status experimental
F;79,82/Modified site: lysine derivative (Lys) (probably 5-hydroxylysine) #status experimental

Alignment Scores:

Pred. No.:	6.47e-55	Length:	238
Score:	705.50	Matches:	143
Percent Similarity:	70.08%	Conservative:	28
Best Local Similarity:	58.61%	Mismatches:	60
Query Match:	51.68%	Indels:	13
DB:	1	Gaps:	4

US-10-054-536-2 (1-747) x LNRTWA (1-238)

QY	19	CTCCCTCTCCTT---CTCCGTAGTATGTTGGCAGCGTCTTACTCAGAACTGTGACCTGT	75
Db	4	LeuProLeuLeuValLeuLeuCysValValSerSerSerGlySerGlnThrCys	23
QY	76	GAGGATGCCCAAGAACCTGCCCTGCAGTGCATTCCTCTGATTCCTCAGGCATCAACGGC	135
Db	24	GluGluThrLeuLysThrCys---SerValIleAlaCys-----	35
QY	136	TTCCAGGCAAGATGGCGTGTATGGCCACCAAGGGAGAGGGGGAACCCAGGCCAAGGG	195
Db	36	-----GlyArgAspGlyArgAspGlyProLysGlyGluLysGlyGluProGlyGlnGly	53
QY	196	CTCAGAGGCTTACAGGCGCCCTCGAAAGTGTGGGCTCCAGGAATCCAGGCGCTCT	255
Db	54	LeuArgGlyLeuGlnGlyProProGlyLysLeuGlyProProGlySerValGlyAlaPro	73
QY	256	GGGTCCACAGGACCAAGGCCCAAGAGGACCCCTGGAAAAAGTCGGATGTGTAGT	315
Db	74	GlySerGlnGlyProLysGlyGlnLysGlyAspArgGlyAspSerArgAlaIleGluVal	93
QY	316	AGCTGTGCT-----GCCTCAGAAAGAAAGCTCTGCAACAGAAATGGCAGCTATCAA	369
Db	94	LysLeuAlaAsnMetGluAlaGluIleAsnThrLeuLysSerLysLeuGluLeuThrAsn	113
QY	370	AGTGGCTGACCTTCTCTGGGCAACAACTTGGGAACAAGTCTTCCTGACCAATGGT	429
Db	114	LysLeuHisAlaPheSerMetGlyLysSerGlyLysPhePheValThrAsnHis	133
QY	430	GAATAATAGCTTTGAAAAAGTAGGCGCTTGTGTCAAGTTCCAGGCTCTGTGGCC	489
Db	134	GluArgMetProPheSerLysValLysAlaLeuCysSerGluLeuArgGlyThrValAla	153

QY	490	ACCCCGAGGAATGCTGCAGAGAAATGCGCAATTCAGAAATCTCATCAAGGAGGAAGCTTC	549
Db	154	IleProArgAsnAlaGluGluAsnLysAlaIleGlnGluValAlaLysThrSerAlaPhe	173
QY	550	CTGGGCATCCTGATGAGAAGACAGAGGCGAGTTTGTGGATCTGACAGGAATAGACTG	609
Db	174	LeuGlyIleThrAspGluValThrGluGlyGlnPheMetTyrValThrGlyArgLeu	193
QY	610	ACCTACACAACTGAAACGAGGGTGAACCCACAAATGCTGTTCTGATGAAGATTGTGA	669
Db	194	ThrTyrSerAsnTrpLysAspGluProAsnAspHisGlySerGlyGluAspCysVal	213
QY	670	TTGCTACTGAAATATGGCCAGTGAATGACGTCCCTCTCCCACTCCCATCTGGCCGTC	729
Db	214	ThrIleValAspAsnGlyLeuTrpAsnAspIleSerCysGlnAlaSerHisThrAlaVal	233
QY	730	TCTGAGTTCCCT 741	
Db	234	CysGluPhePro 237	

RESULT 6
A45225
N;Alternate names: SP-D
C;Species: Homo sapiens (man)
C;Date: 16-Apr-1999 #sequence revision 16-Apr-1999 #text change 22-Jun-1999
C;Accession: A45225; S23434; S24555; S44420; S18382; A56776
R;Crouch, E.; Rust, K.; Veille, R.; Donis-Keller, H.; Grosso, L.
J. Biol. Chem. 268, 2976-2983, 1993
A;Title: Genomic organization of human surfactant protein D (SP-D). SP-D is encoded on chromosome 10p15
A;Reference number: A45225; MUID:93155122; PMID:8428971
A;Accession: A45225
A;Molecule type: DNA
A;Residues: 1-375 <CRO>
A;Cross-references: GB:L05483; GB:L05484; GB:L05485; NID:g292505; PIDN:AA859450.1; PID:g292505
A;Experimental source: placenta
A;Note: sequence extracted from NCBI backbone (NCBIP:124316)
R;Lu, J.; Willis, A.C.; Reid, K.B.M.
Biochem. J. 284, 795-802, 1992
A;Title: Purification, characterization and cDNA cloning of human lung surfactant protein D
A;Reference number: S23434; MUID:92322003; PMID:1339284
A;Accession: S23434
A;Molecule type: mRNA
A;Residues: 1-30, 'T', 32-121, 'P', 123-179, 'A', 181-375 <LUJ1>
A;Cross-references: EMBL:X65018; NID:g34766; PIDN:CAA46152.1; PID:g34767
A;Experimental source: lung
A;Accession: S24555
A;Molecule type: protein
A;Residues: 214-234, 'X', 236, 'XX', 239-241 <LUJ2>
R;Hoppe, H.J.; Barlow, P.N.; Reid, K.B.M.
FEBS Lett. 344, 191-195, 1994
A;Title: A parallel three stranded alpha-helical bundle at the nucleation site of collagen
A;Reference number: S44420; MUID:94244769; PMID:8187882
A;Accession: S44420
A;Molecule type: mRNA
A;Residues: 202-257 <HOP>
R;Rust, K.; Grosso, L.; Zhang, V.; Chang, D.; Persson, A.; Longmore, W.; Cai, G.Z.; Crou
Arch. Biochem. Biophys. 290, 116-126, 1991
A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recognition domain
A;Reference number: S18382; MUID:91378578; PMID:1898081
A;Accession: S18382
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 'F', 60-205, 'P', 207-374, 'HF', <RUS>
A;Cross-references: GB:L05485; NID:g292505
A;Note: corrections to this sequence are reported in reference A56776
R;Crouch, E.; Persson, A.; Chang, D.
Am. J. Pathol. 142, 241-248, 1993
A;Title: Accumulation of surfactant protein D in human pulmonary alveolar proteinosis.
A;Reference number: A56776; MUID:93142849; PMID:8424457
A;Accession: A56776
A;Status: preliminary

A:Molecule type: protein
A:Residues: 46-58,'F','60-62','E','64-72;223-227','X','229-239','P','241-245','X','247-256','X','258-260','X',261-262,'X',263-264,'X',265-266,'X',267-268,'X',269-270,'X',271-272,'X',273-274,'X',275-276,'X',277-278,'X',279-280,'X',281-282,'X',283-284,'X',285-286,'X',287-288,'X',289-290,'X',291-292,'X',293-294,'X',295-296,'X',297-298,'X',299-300,'X',301-302,'X',303-304,'X',305-306,'X',307-308,'X',309-310,'X',311-312,'X',313-314,'X',315-316,'X',317-318,'X',319-320,'X',321-322,'X',323-324,'X',325-326,'X',327-328,'X',329-330,'X',331-332,'X',333-334,'X',335-336,'X',337-338,'X',339-340,'X',341-342,'X',343-344,'X',345-346,'X',347-348,'X',349-350,'X',351-352,'X',353-354,'X',355-356,'X',357-358,'X',359-360,'X',361-362,'X',363-364,'X',365-366,'X',367-368,'X',369-370,'X',371-372,'X',373-374,'X',375-376,'X',377-378,'X',379-380,'X',381-382,'X',383-384,'X',385-386,'X',387-388,'X',389-390,'X',391-392,'X',393-394,'X',395-396,'X',397-398,'X',399-400,'X',401-402,'X',403-404,'X',405-406,'X',407-408,'X',409-410,'X',411-412,'X',413-414,'X',415-416,'X',417-418,'X',419-420,'X',421-422,'X',423-424,'X',425-426,'X',427-428,'X',429-430,'X',431-432,'X',433-434,'X',435-436,'X',437-438,'X',439-440,'X',441-442,'X',443-444,'X',445-446,'X',447-448,'X',449-450,'X',451-452,'X',453-454,'X',455-456,'X',457-458,'X',459-460,'X',461-462,'X',463-464,'X',465-466,'X',467-468,'X',469-470,'X',471-472,'X',473-474,'X',475-476,'X',477-478,'X',479-480,'X',481-482,'X',483-484,'X',485-486,'X',487-488,'X',489-490,'X',491-492,'X',493-494,'X',495-496,'X',497-498,'X',499-500,'X',501-502,'X',503-504,'X',505-506,'X',507-508,'X',509-510,'X',511-512,'X',513-514,'X',515-516,'X',517-518,'X',519-520,'X',521-522,'X',523-524,'X',525-526,'X',527-528,'X',529-530,'X',531-532,'X',533-534,'X',535-536,'X',537-538,'X',539-540,'X',541-542,'X',543-544,'X',545-546,'X',547-548,'X',549-550,'X',551-552,'X',553-554,'X',555-556,'X',557-558,'X',559-560,'X',561-562,'X',563-564,'X',565-566,'X',567-568,'X',569-570,'X',571-572,'X',573-574,'X',575-576,'X',577-578,'X',579-580,'X',581-582,'X',583-584,'X',585-586,'X',587-588,'X',589-590,'X',591-592,'X',593-594,'X',595-596,'X',597-598,'X',599-600,'X',601-602,'X',603-604,'X',605-606,'X',607-608,'X',609-610,'X',611-612,'X',613-614,'X',615-616,'X',617-618,'X',619-620,'X',621-622,'X',623-624,'X',625-626,'X',627-628,'X',629-630,'X',631-632,'X',633-634,'X',635-636,'X',637-638,'X',639-640,'X',641-642,'X',643-644,'X',645-646,'X',647-648,'X',649-650,'X',651-652,'X',653-654,'X',655-656,'X',657-658,'X',659-660,'X',661-662,'X',663-664,'X',665-666,'X',667-668,'X',669-670,'X',671-672,'X',673-674,'X',675-676,'X',677-678,'X',679-680,'X',681-682,'X',683-684,'X',685-686,'X',687-688,'X',689-690,'X',691-692,'X',693-694,'X',695-696,'X',697-698,'X',699-700,'X',701-702,'X',703-704,'X',705-706,'X',707-708,'X',709-710,'X',711-712,'X',713-714,'X',715-716,'X',717-718,'X',719-720,'X',721-722,'X',723-724,'X',725-726,'X',727-728,'X',729-730,'X',731-732,'X',733-734,'X',735-736,'X',737-738,'X',739-740,'X',741-742,'X',743-744,'X',745-746,'X',747-748,'X',749-750,'X',751-752,'X',753-754,'X',755-756,'X',757-758,'X',759-760,'X',761-762,'X',763-764,'X',765-766,'X',767-768,'X',769-770,'X',771-772,'X',773-774,'X',775-776,'X',777-778,'X',779-780,'X',781-782,'X',783-784,'X',785-786,'X',787-788,'X',789-790,'X',791-792,'X',793-794,'X',795-796,'X',797-798,'X',799-800,'X',801-802,'X',803-804,'X',805-806,'X',807-808,'X',809-810,'X',811-812,'X',813-814,'X',815-816,'X',817-818,'X',819-820,'X',821-822,'X',823-824,'X',825-826,'X',827-828,'X',829-830,'X',831-832,'X',833-834,'X',835-836,'X',837-838,'X',839-840,'X',841-842,'X',843-844,'X',845-846,'X',847-848,'X',849-850,'X',851-852,'X',853-854,'X',855-856,'X',857-858,'X',859-860,'X',861-862,'X',863-864,'X',865-866,'X',867-868,'X',869-870,'X',871-872,'X',873-874,'X',875-876,'X',877-878,'X',879-880,'X',881-882,'X',883-884,'X',885-886,'X',887-888,'X',889-890,'X',891-892,'X',893-894,'X',895-896,'X',897-898,'X',899-900,'X',901-902,'X',903-904,'X',905-906,'X',907-908,'X',909-910,'X',911-912,'X',913-914,'X',915-916,'X',917-918,'X',919-920,'X',921-922,'X',923-924,'X',925-926,'X',927-928,'X',929-930,'X',931-932,'X',933-934,'X',935-936,'X',937-938,'X',939-940,'X',941-942,'X',943-944,'X',945-946,'X',947-948,'X',949-950,'X',951-952,'X',953-954,'X',955-956,'X',957-958,'X',959-960,'X',961-962,'X',963-964,'X',965-966,'X',967-968,'X',969-970,'X',971-972,'X',973-974,'X',975-976,'X',977-978,'X',979-980,'X',981-982,'X',983-984,'X',985-986,'X',987-988,'X',989-990,'X',991-992,'X',993-994,'X',995-996,'X',997-998,'X',999-1000,'X',1001-1002,'X',1003-1004,'X',1005-1006,'X',1007-1008,'X',1009-1010,'X',1011-1012,'X',1013-1014,'X',1015-1016,'X',1017-1018,'X',1019-1020,'X',1021-1022,'X',1023-1024,'X',1025-1026,'X',1027-1028,'X',1029-1030,'X',1031-1032,'X',1033-1034,'X',1035-1036,'X',1037-1038,'X',1039-1040,'X',1041-1042,'X',1043-1044,'X',1045-1046,'X',1047-1048,'X',1049-1050,'X',1051-1052,'X',1053-1054,'X',1055-1056,'X',1057-1058,'X',1059-1060,'X',1061-1062,'X',1063-1064,'X',1065-1066,'X',1067-1068,'X',1069-1070,'X',1071-1072,'X',1073-1074,'X',1075-1076,'X',1077-1078,'X',1079-1080,'X',1081-1082,'X',1083-1084,'X',1085-1086,'X',1087-1088,'X',1089-1090,'X',1091-1092,'X',1093-1094,'X',1095-1096,'X',1097-1098,'X',1099-1100,'X',1101-1102,'X',1103-1104,'X',1105-1106,'X',1107-1108,'X',1109-1110,'X',1111-1112,'X',1113-1114,'X',1115-1116,'X',1117-1118,'X',1119-1120,'X',1121-1122,'X',1123-1124,'X',1125-1126,'X',1127-1128,'X',1129-1130,'X',1131-1132,'X',1133-1134,'X',1135-1136,'X',1137-1138,'X',1139-1140,'X',1141-1142,'X',1143-1144,'X',1145-1146,'X',1147-1148,'X',1149-1150,'X',1151-1152,'X',1153-1154,'X',1155-1156,'X',1157-1158,'X',1159-1160,'X',1161-1162,'X',1163-1164,'X',1165-1166,'X',1167-1168,'X',1169-1170,'X',1171-1172,'X',1173-1174,'X',1175-1176,'X',1177-1178,'X',1179-1180,'X',1181-1182,'X',1183-1184,'X',1185-1186,'X',1187-1188,'X',1189-1190,'X',1191-1192,'X',1193-1194,'X',1195-1196,'X',1197-1198,'X',1199-1200,'X',1201-1202,'X',1203-1204,'X',1205-1206,'X',1207-1208,'X',1209-1210,'X',1211-1212,'X',1213-1214,'X',1215-1216,'X',1217-1218,'X',1219-1220,'X',1221-1222,'X',1223-1224,'X',1225-1226,'X',1227-1228,'X',1229-1230,'X',1231-1232,'X',1233-1234,'X',1235-1236,'X',1237-1238,'X',1239-1240,'X',1241-1242,'X',1243-1244,'X',1245-1246,'X',1247-1248,'X',1249-1250,'X',1251-1252,'X',1253-1254,'X',1255-1256,'X',1257-1258,'X',1259-1260,'X',1261-1262,'X',1263-1264,'X',1265-1266,'X',1267-1268,'X',1269-1270,'X',1271-1272,'X',1273-1274,'X',1275-1276,'X',1277-1278,'X',1279-1280,'X',1281-1282,'X',1283-1284,'X',1285-1286,'X',1287-1288,'X',1289-1290,'X',1291-1292,'X',1293-1294,'X',1295-1296,'X',1297-1298,'X',1299-1300,'X',1301-1302,'X',1303-1304,'X',1305-1306,'X',1307-1308,'X',1309-1310,'X',1311-1312,'X',1313-1314,'X',1315-1316,'X',1317-1318,'X',1319-1320,'X',1321-1322,'X',1323-1324,'X',1325-1326,'X',1327-1328,'X',1329-1330,'X',1331-1332,'X',1333-1334,'X',1335-1336,'X',1337-1338,'X',1339-1340,'X',1341-1342,'X',1343-1344,'X',1345-1346,'X',1347-1348,'X',1349-1350,'X',1351-1352,'X',1353-1354,'X',1355-1356,'X',1357-1358,'X',1359-1360,'X',1361-1362,'X',1363-1364,'X',1365-1366,'X',1367-1368,'X',1369-1370,'X',1371-1372,'X',1373-1374,'X',1375-1376,'X',1377-1378,'X',1379-1380,'X',1381-1382,'X',1383-1384,'X',1385-1386,'X',1387-1388,'X',1389-1390,'X',1391-1392,'X',1393-1394,'X',1395-1396,'X',1397-1398,'X',1399-1400,'X',1401-1402,'X',1403-1404,'X',1405-1406,'X',1407-1408,'X',1409-1410,'X',1411-1412,'X',1413-1414,'X',1415-1416,'X',1417-1418,'X',1419-1420,'X',1421-1422,'X',1423-1424,'X',1425-1426,'X',1427-1428,'X',1429-1430,'X',1431-1432,'X',1433-1434,'X',1435-1436,'X',1437-1438,'X',1439-1440,'X',1441-1442,'X',1443-1444,'X',1445-1446,'X',1447-1448,'X',1449-1450,'X',1451-1452,'X',1453-1454,'X',1455-1456,'X',1457-1458,'X',1459-1460,'X',1461-1462,'X',1463-1464,'X',1465-1466,'X',1467-1468,'X',1469-1470,'X',1471-1472,'X',1473-1474,'X',1475-1476,'X',1477-1478,'X',1479-1480,'X',1481-1482,'X',1483-1484,'X',1485-1486,'X',1487-1488,'X',1489-1490,'X',1491-1492,'X',1493-1494,'X',1495-1496,'X',1497-1498,'X',1499-1500,'X',1501-1502,'X',1503-1504,'X',1505-1506,'X',1507-1508,'X',1509-1510,'X',1511-1512,'X',1513-1514,'X',1515-1516,'X',1517-1518,'X',1519-1520,'X',1521-1522,'X',1523-1524,'X',1525-1526,'X',1527-1528,'X',1529-1530,'X',1531-1532,'X',1533-1534,'X',1535-1536,'X',1537-1538,'X',1539-1540,'X',1541-1542,'X',1543-1544,'X',1545-1546,'X',1547-1548,'X',1549-1550,'X',1551-1552,'X',1553-1554,'X',1555-1556,'X',1557-1558,'X',1559-1560,'X',1561-1562,'X',1563-1564,'X',1565-1566,'X',1567-1568,'X',1569-1570,'X',1571-1572,'X',1573-1574,'X',1575-1576,'X',1577-1578,'X',1579-1580,'X',1581-1582,'X',1583-1584,'X',1585-1586,'X',1587-1588,'X',1589-1590,'X',1591-1592,'X',1593-1594,'X',1595-1596,'X',1597-1598,'X',1599-1600,'X',1601-1602,'X',1603-1604,'X',1605-1606,'X',1607-1608,'X',1609-1610,'X',1611-1612,'X',1613-1614,'X',1615-1616,'X',1617-1618,'X',1619-1620,'X',1621-1622,'X',1623-1624,'X',1625-1626,'X',1627-1628,'X',1629-1630,'X',1631-1632,'X',1633-1634,'X',1635-1636,'X',1637-1638,'X',1639-1640,'X',1641-1642,'X',1643-1644,'X',1645-1646,'X',1647-1648,'X',1649-1650,'X',1651-1652,'X',1653-1654,'X',1655-1656,'X',1657-1658,'X',1659-1660,'X',1661-1662,'X',1663-1664,'X',1665-1666,'X',1667-1668,'X',1669-1670,'X',1671-1672,'X',1673-1674,'X',1675-1676,'X',1677-1678,'X',1679-1680,'X',1681-1682,'X',1683-1684,'X',1685-1686,'X',1687-1688,'X',1689-1690,'X',1691-1692,'X',1693-1694,'X',1695-1696,'X',1697-1698,'X',1699-1700,'X',1701-1702,'X',1703-1704,'X',1705-1706,'X',1707-1708,'X',1709-1710,'X',1711-1712,'X',1713-1714,'X',1715-1716,'X',1717-1718,'X',1719-1720,'X',1721-1722,'X',1723-1724,'X',1725-1726,'X',1727-1728,'X',1729-1730,'X',1731-1732,'X',1733-1734,'X',1735-1736,'X',1737-1738,'X',1739-1740,'X',1741-1742,'X',1743-1744,'X',1745-1746,'X',1747-1748,'X',1749-1750,'X',1751-1752,'X',1753-1754,'X',1755-1756,'X',1757-1758,'X',1759-1760,'X',1761-1762,'X',1763-1764,'X',1765-1766,'X',1767-1768,'X',1769-1770,'X',1771-1772,'X',1773-1774,'X',1775-1776,'X',1777-1778,'X',1779-1780,'X',1781-1782,'X',1783-1784,'X',1785-1786,'X',1787-1788,'X',1789-1790,'X',1791-1792,'X',1793-1794,'X',1795-1796,'X',1797-1798,'X',1799-1800,'X',1801-1802,'X',1803-1804,'X',1805-1806,'X',1807-1808,'X',1809-1810,'X',1811-1812,'X',1813-1814,'X',1815-1816,'X',1817-1818,'X',1819-1820,'X',1821-1822,'X',1823-1824,'X',1825-1826,'X',1827-1828,'X',1829-1830,'X',1831-1832,'X',1833-1834,'X',1835-1836,'X',1837-1838,'X',1839-1840,'X',1841-1842,'X',1843-1844,'X',1845-1846,'X',1847-1848,'X',1849-1850,'X',1851-1852,'X',1853-1854,'X',1855-1856,'X',1857-1858,'X',1859-1860,'X',1861-1862,'X',1863-1864,'X',1865-1866,'X',1867-1868,'X',1869-1870,'X',1871-1872,'X',1873-1874,'X',1875-1876,'X',1877-1878,'X',1879-1880,'X',1881-1882,'X',1883-1884,'X',1885-1886,'X',1887-1888,'X',1889-1890,'X',1891-1892,'X',1893-1894,'X',1895-1896,'X',1897-1898,'X',1899-1900,'X',1901-1902,'X',1903-1904,'X',1905-1906,'X',1907-1908,'X',1909-1910,'X',1911-1912,'X',1913-1914,'X',1915-1916,'X',1917-1918,'X',1919-1920,'X',1921-1922,'X',1923-1924,'X',1925-1926,'X',1927-1928,'X',1929-1930,'X',1931-1932,'X',1933-1934,'X',1935-1936,'X',1937-1938,'X',1939-1940,'X',1941-1942,'X',1943-1944,'X',1945-1946,'X',1947-1948,'X',1949-1950,'X',1951-1952,'X',1953-1954,'X',1955-1956,'X',1957-1958,'X',1959-1960,'X',1961-1962,'X',1963-1964,'X',1965-1966,'X',1967-1968,'X',1969-1970,'X',1971-1972,'X',1973-1974,'X',1975-1976,'X',1977-1978,'X',1979-1980,'X',1981-1982,'X',1983-1984,'X',1985-1986,'X',1987-1988,'X',1989-1990,'X',1991-1992,'X',1993-1994,'X',1995-1996,'X',1997-1998,'X',1999-2000,'X',2001-2002,'X',2003-2004,'X',2005-2006,'X',2007-2008,'X',2009-2010,'X',2011-2012,'X',2013-2014,'X',2015-2016,'X',2017-2018,'X',2019-2020,'X',2021-2022,'X',2023-2024,'X',2025-2026,'X',2027-2028,'X',2029-2030,'X',2031-2032,'X',2033-2034,'X',2035-2036,'X',2037-2038,'X',2039-2040,'X',2041-2042,'X',2043-2044,'X',2045-2046,'X',2047-2048,'X',2049-2050,'X',2051-2052,'X',2053-2054,'X',2055-2056,'X',2057-2058,'X',2059-2060,'X',2061-2062,'X',2063-2064,'X',2065-2066,'X',2067-2068,'X',2069-2070,'X',2071-2072,'X',2073-2074,'X',2075-2076,'X',2077-2078,'X',2079-2080,'X',2081-2082,'X',2083-2084,'X',2085-2086,'X',2087-2088,'X',2089-2090,'X',2091-2092,'X',2093-2094,'X',2095-2096,'X',2097-2098,'X',2099-2100,'X',2101-2102,'X',2103-2104,'X',2105-2106,'X',2107-2108,'X',2109-2110,'X',2111-2112,'X',2113-2114,'X',2115-2116,'X',2117-2118,'X',2119-2120,'X',2121-2122,'X',2123-2124,'X',2125-2126,'X',2127-2128,'X',2129-2130,'X',2131-2132,'X',2133-2134,'X',2135-2136,'X',2137-2138,'X',2139-2140,'X',2141-2142,'X',2143-2144,'X',2145-2146,'X',2147-2148,'X',2149-2150,'X',2151-2152,'X',2153-2154,'X',2155-2156,'X',2157-2158,'X',2159-2160,'X',2161-2162,'X',2163-2164,'X',2165-2166,'X',2167-2168,'X',2169-2170,'X',2171-2172,'X',2173-2174,'X',2175-2176,'X',2177-2178,'X',2179-2180,'X',2181-2182,'X',2183-2184,'X',2185-2186,'X',2187-2188,'X',2189-2190,'X',2191-2192,'X',2193-2194,'X',2195-2196,'X',2197-2198,'X',2199-2200,'X',2201-2202,'X',2203-2204,'X',2205-2206,'X',2207-2208,'X',2209-2210,'X',2211-2212,'X',2213-2214,'X',2215-2216,'X',2217-2218,'X',2219-2220,'X',2221-2222,'X',2223-2224,'X',2225-2226,'X',2227-2228,'X',2229-2230,'X',2231-2232,'X',2233-2234,'X',2235-2236,'X',2237-2238,'X',2239-2240,'X',2241-2242,'X',2243-2244,'X',2245-2246,'X',2247-2248,'X',2249-2250,'X',2251-2252,'X',2253-2254,'X',2255-2256,'X',2257-2258,'X',2259-2260,'X',2261-2262,'X',2263-2264,'X',2265-2266,'X',2267-2268,'X',2269-2270,'X',2271-2272,'X',2273-2274,'X',2275-2276,'X',2277-2278,'X',2279-2280,'X',2281-2282,'X',2283-2284,'X',2285-2286,'X',2287-2288,'X',2289-2290,'X',2291-2292,'X',2293-2294,'X',2295-2296,'X',2297-2298,'X',2299-2300,'X',2301-2302,'X',2303-2304,'X',2305-2306,'X',2307-2308,'X',2309-2310,'X',2311-2312,'X',2313-2314,'X',2315-2316,'X',2317-2318,'X',2319-2320,'X',2321-2322,'X',2323-2324,'X',2325-2326,'X',2327-2328,'X',2329-2330,'X',2331-2332,'X',2333-2334,'X',2335-2336,'X',2337-2338,'X',2339-2340,'X',2341-2342,'X',2343-2344,'X',2345-2346,'X',2347-2348,'X',2349-2350,'X',2351-2352,'X',2353-2354,'X',2355-2356,'X',2357-2358,'X',2359-2360,'X',2361-2362,'X',2363-2364,'X',2365-2366,'X',2367-2368,'X',2369-2370,'X',2371-2372,'X',2373-2374,'X',2375-2376,'X',2377-2378,'X',2379-2380,'X',2381-2382,'X',2383-2384,'X',2385-2386,'X',2387-2388,'X',2389-2390,'X',2391-2392,'X',2393-2394,'X',2395-2396,'X',2397-2398,'X',2399-2400,'X',2401-2402,'X',2403-2404,'X',2405-2406,'X',2407-2408,'

QY 309 ----- 309
 Db LysGlyGluArgGlyAlaProGlyGluGlnGlyAlaProGlyAsnAlaGlyAlaGly 183
 QY 309 ----- 309
 Db ProAlaGlyProAlaGlyProGlnGlyAlaProGlySerArgGlyProProGlyLeuLys 203
 QY 310 -----CATAGT 315
 Db 204 GlyAspArgGlyAlaProGlyAspArgGlyLysGlyGluSerGlyLeuProAspSer 223
 QY 316 AGCCTGGCTGCTCAGAAAGAAAGCTTCGAAACAGAAATGGCAGCTATC ----- 366
 Db 224 AlaAlaLeuArgGlnGlnMetGluAlaLeuAsnGlyLysLeuGlnArgLeuGluAla 243
 QY 367 -----AAAAATGGCTGACCTTCTCTCGGCCAACAAAGTTGGGAACAGTTTC 414
 Db 244 PheSerArgTyrLysLysAlaAlaLeuPheProAspGlyGlnSerValGlyAspLysIle 263
 QY 415 TTCCTGACCAATGGTGAATAATGACCTTTCGAAAGTGAAGCCCTTGTGTCAAGTTC 474
 Db 264 PheArgAlaAlaAsnSerGluGluProPheGluAspAlaLysGluMetCysArgGlnAla 283
 QY 475 CAGGCCTCTGTGGCCACCCAGGAATGCTGCAGAGAATGGAGCCATTCAGAATCTCATC 534
 Db 284 GlyGlyGlnLeuAlaSerProArgSerAlaThrGluAsnAlaAlaValGlnGlnLeuVal 303
 QY 535 -----AAGGAGGAAGCCTTCTCTGGGCATCACTGATGAGACAGAGGGGAGTTT 585
 Db 304 ThrAlaHisSerLysAlaAlaPheLeuSerMetThrAspValGlyThrGluGlyLysPhe 323
 QY 586 GTGCATCTGACAGAAATAGACTGACCTACACAACTGGAAGGAGTGAACCCACAAAT 645
 Db 324 ThrTyrProThrGlyGluAlaLeuValTyrSerAsnTrpAlaProGlyGluProAsnAsn 343
 QY 646 GCTGCTTGTGATGAAGATTGTGTTACTGAAAAATGCCAGTGAATGAGTCCCTCC 705
 Db 344 AsnGlyGlyAlaGluAsnCysValGluIlePheThrAsnGlyGlnTrpAsnAspLysAla 363
 QY 706 TGCTCCACCTCCATCTGGCCGTCTGTGAGTTC 738
 Db 364 CysGlyGluGlnArgLeuValIleCysGluPhe 374

RESULT 9

I45878
 collectin-43 - bovine
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 20-Aug-1999
 C;Accession: I45878
 R;Liu, L.S.; Sastry, R.; Hartshorn, K.L.; Lee, Y.M.; Okarma, T.B.; Tauber, A.I.; Sastry
 Gene 141, 277-281, 1994
 A;Title: Bovine collectin (BC) mRNA expressed in liver: cloning and characterization of
 A;Reference number: I45878; MUID:94215917; PMID:8163202
 A;Accession: I45878
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-371 <L10>
 A;Cross-references: GB:L18871; NID:9495012; PIDN:AAA20126.1; PID:9495013
 C;Superfamily: pulmonary surfactant protein D; C-type lectin homology
 F;248-369/Domain: C-type lectin homology <LCH>

Alignment Scores:

Pred. No.:	3-35e-31	Length:	371
Score:	440.50	Matches:	91
Percent Similarity:	56.77%	Conservative:	39
Best Local Similarity:	39.74%	Mismatches:	74
Query Match:	32.27%	Indels:	25
DB:	2	Gaps:	6

US-10-054-536-2 (1-747) x I45878 (1-371)

QY 118 TCTCCAGGCATCAACGGCTTCCAGGCAAGATGGCGTGTATGGCACCAAGGAGAAAG 177
 Db 146 AlaProGlyIleGlnGlyPheProGlyPro-----SerGlyLeuLysGlyGluLys 162
 QY 178 GGGGAACACAGGCCAA-----GGGCTCAGAGGCTTACAGGGCCCCCTCGAAAG 225
 Db 163 GlyAlaProGlyGluThrGlyAlaProGlyHisAlaGlyValThrGlyProSerGlyAla 182
 QY 226 TTGGGCGCTCCAGGAATCCAGGGCTTCTGGGTCCAGGACCAAGGCGGCAAAAGGA 285
 Db 183 IleGlyProGlnGlyProSerGlyAlaArgGlyProProGlyLeuLysGlyAspArgGly 202
 QY 286 GACCTCGAAAAGT---CCGATGGTGTATAGTAGCCCTGGCTGCC----- 327
 Db 203 AspProGlyGluThrGlyAlaLysGlyGluSerGlyLeuAlaGluValAsnAlaLeuLys 222
 QY 328 -----TCAGAAAGAAAGCTCTGCAACAGAAATGGCAGTATC 366
 Db 223 GlnArgValThrIleLeuAspGlyHisLeuArgPheGlnAsnAlaPheSerGlnTyr 242
 QY 367 AAAAAGTGGCTGACCTTCTCTGGGCAACAAAGTTGGGAACAAAGTTCTCTGACCAAT 426
 Db 243 LysLysAlaValLeuPheProAspGlyGlnAlaValGlyGluLysIlePheLysThrAla 262
 QY 427 GGTGAAATAATGACCTTTGAAAAAGTGAAGCCTTGTGTGTCAGTTCAGGCTCTCTGTG 486
 Db 263 GlyAlaValLysSerTyrSerAspAlaGluGlnLeuCysArgGluAlaLysGlyGlnLeu 282
 QY 487 GCCACCCCGAGGAATGCTGCAGAGAATGGAGCCATTCAGAATCTCATCAAG----- 537
 Db 283 AlaSerProArgSerSerAlaGluAsnGluAlaValThrGlnMetValArgAlaGlnGlu 302
 QY 538 GAGGAAGCCTCTCTGGGCATCCTGATGAGAGACAGAGGCGAGTTGTGTGATCTGACA 597
 Db 303 LysAsnAlaTyrLeuSerMetAsnAspIleSerThrGluGlyArgPheThrTyrProThr 322
 QY 598 GGAATAGACTGACCTACAAACTGGAACGAGGCTGAACCCCAACAATGCT-----GGT 651
 Db 323 GlyGluLeuLeuValTyrSerAsnTrpAlaAspGlyGluProAsnAsnSerAspGluGly 342
 QY 652 TCTGATGAAGATTGTGTTACTGAAATAATGCCAGTGAATGAGTCCCTCTGCTCC 711
 Db 343 GlnProGluAsnCysValGluIlePheProAspGlyLysTrpAsnAspValProCysSer 362
 QY 712 ACCTCCCATCTGGCGTCTGTGAGTTC 738
 Db 363 LysGlnLeuLeuValIleCysGluPhe 371

RESULT 10

A53570
 collectin-43 - bovine
 N;Alternate names: lectin CL-43
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 12-Apr-1995 #sequence_revision 23-Feb-1996 #text_change 17-Mar-1999
 C;Accession: A53570; A46689
 R;Lim, B.L.; Willis, A.C.; Reid, K.B.M.; Lu, J.; Laursen, S.B.; Jensenius, J.C.; Holmskov
 J. Biol. Chem. 269, 11820-11824, 1994
 A;Title: Primary structure of bovine collectin-43 (CL-43). Comparison with conglutinin
 A;Reference number: A53570; MUID:94216283; PMID:8163480
 A;Accession: A53570
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-301 <LIM>
 A;Cross-references: GB:X75912
 R;Holmskov, U.; Teisener, B.; Willis, A.C.; Reid, K.B.; Jensenius, J.C.
 J. Biol. Chem. 268, 10120-10125, 1993
 A;Title: Purification and characterization of a bovine serum lectin (CL-43) with structure
 A;Reference number: A46689; MUID:93252891; PMID:8486682
 A;Accession: A46689
 A;Molecule type: protein
 A;Residues: 1-27 <HOL>
 A;Experimental source: serum
 A;Note: sequence extracted from NCBI backbone (NCBIP:131234)

Result No.	Score	Query		Length	PB	ID	Description
		Match	%				
1	747	100.0	900	4	US-09-198-603C-1		Sequence 1, Appli
2	204.4	27.4	1211	4	US-09-198-603C-25		Sequence 25, Appl
3	126	16.9	714	4	US-09-198-603C-26		Sequence 26, Appl
4	73.4	9.8	1558	4	US-09-198-603C-24		Sequence 24, Appl
5	67.6	9.0	885	1	US-08-365-103B-3		Sequence 3, Appl
6	67.6	9.0	924	1	US-08-365-103B-5		Sequence 5, Appl
7	67.6	9.0	1005	1	US-08-365-103B-1		Sequence 1, Appli
8	62.8	8.4	369	4	US-09-535-521-24		Sequence 24, Appl
C 9	62.8	8.4	369	4	US-09-535-521-26		Sequence 26, Appl
10	62.8	8.4	384	4	US-09-535-521-7		Sequence 7, Appl
C 11	62.8	8.4	384	4	US-09-535-521-9		Sequence 9, Appli
12	62.8	8.4	417	4	US-09-535-521-10		Sequence 10, Appl
C 13	62.8	8.4	417	4	US-09-535-521-12		Sequence 12, Appl
14	62.8	8.4	423	4	US-09-535-521-13		Sequence 13, Appl
C 15	62.8	8.4	423	4	US-09-535-521-15		Sequence 15, Appl
16	62.8	8.4	561	4	US-09-535-521-16		Sequence 16, Appl
C 17	62.8	8.4	561	4	US-09-535-521-18		Sequence 18, Appl
18	62.8	8.4	624	4	US-09-535-521-19		Sequence 19, Appl
C 19	62.8	8.4	624	4	US-09-535-521-21		Sequence 21, Appl
20	62.8	8.4	876	4	US-09-535-521-4		Sequence 4, Appli
C 21	62.8	8.4	876	4	US-09-535-521-6		Sequence 6, Appli
22	62.8	8.4	2851	4	US-09-535-521-1		Sequence 1, Appli
C 23	62.8	8.4	2851	4	US-09-535-521-3		Sequence 3, Appli
C 24	55.2	7.4	7218	1	US-08-232-463-14		Sequence 14, Appl
25	52.4	7.0	1152	4	US-09-776-976-1		Sequence 1, Appli
26	52.4	7.0	1152	4	US-09-909-547-1		Sequence 1, Appli
27	51.4	6.9	1276	2	US-08-463-911-1		Sequence 1, Appli


```
QY 601 AATAGACTGACCTACAAAGTGAACGAGGGTGAACCCACAAATGCTGTTCTGATGAA 660
Db 577 GGGCCITTAACCTACGCAACTGGAAACCTGGAGAACCAATAATACAAA---ATGAA 633
QY 661 GATTGTGTATTGCTACTGAAATAATGGCCAGTGAATGACGCTCCCTGCTCCACCTCCCAT 720
Db 634 GACTGTGGGTGATAGAGACTCTGGAAATGGAATGATTAGACTGTTCAAATTCAAAT 693
QY 721 CT 722
Db 694 AT 695

RESULT 4
US-09-198-603C-24
; Sequence 24, Application US/09198603C
; Patent No. 6337193
; GENERAL INFORMATION:
; APPLICANT: TULLY, Raymond E.
; APPLICANT: CAUTAGIRONE, G. Thomas
; APPLICANT: MOYER, Shawn S.
; APPLICANT: RONNING, Michael T.
; TITLE OF INVENTION: EXPRESSION OF MANNOSE-BINDING PROTEIN IN METHYLOTROPHIC
; TITLE OF INVENTION: YEAST
; FILE REFERENCE: A7290
; CURRENT APPLICATION NUMBER: US/09/198,603C
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 24
; LENGTH: 1558
; TYPE: DNA
; ORGANISM: RAT
US-09-198-603C-24

Query Match 9.8%; Score 73.4; DB 4; Length 1558;
Best Local Similarity 77.4%; Pred. No. 3.5e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 185 CAGGCCAAGGGCTCAGAGGCTTACAGGGCCCCCTGGAAAGTTGGGGCTCCAGAAATC 244
Db 1173 CAGGTCAGGGCTCAGGGGCTTTCAGAGGCTTCCAGGGAACTGGGGCTCCAGGAAGTG 1232
QY 245 CAGGGCTTCTGGGTACAGGACCAAGGGCCAAAGGACGACCCCTGGAAGAAG 299
Db 1233 TAGGAGCCCTCGAAGTCAAGGACCAAGGCAAGGCAAGGAGGATCGTGAGACAG 1287

RESULT 5
US-08-365-103B-3
; Sequence 3, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G.
; APPLICANT: Nunez, Raphael D.
; APPLICANT: Yodoi, Jungi
; TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,103B
```

```
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S. 37,719
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: Ulif N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 24...884
US-08-365-103B-3

Query Match 9.0%; Score 67.6; DB 1; Length 885;
Best Local Similarity 55.6%; Pred. No. 1.8e-11;
Matches 130; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 484 GTGGCCACCCCGAATGCTGCAGAGATGGAGCCATTGAGATCTCATCAAGGAGAA 543
Db 567 GTCAGCATCCACAGCAAAAGGACAGGACTTCTGTATGCAACATCAACAGAGGAT 626
QY 544 GCCTTCCTGGGCATCACTGATGAGAGAGACAGAAAGGCGCAGTTTGTGATCTGACAGGAAAT 603
Db 627 TCCTGGATTGGCTCCAGGATCTCAATATGGAGGAGAGTTGTATGGTCGACGGGAGC 686
QY 604 AGACTGACCTACAAAACCTGGAACGAGGGTGAAACCCAAACATGCTGTTCTGATGAAGT 563
Db 687 CCTGTGGGTTATAGCAACTGGAATCCAGGGGAGGCCAATAACGGGGGCCAGGGTGAGGAC 746
QY 664 TGTGTATTGCTACTGAAAATAATGGCCAGTGGATGACGTCCTCTCCACCTCC 717
Db 747 TGTGTGATGATGCGGGGATCCGGCCAGTGGAAACGACGCCCTTCTGCCGACGTAC 800

RESULT 6
US-08-365-103B-5
; Sequence 5, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G.
; APPLICANT: Nunez, Raphael D.
; APPLICANT: Yodoi, Jungi
; TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
```

RESULT 7
US-08-365-103B-1
; Sequence 1, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; APPLICANT: Yodoi, Junji
; TITLE OF INVENTION: DNA Sequences for Soluble Proms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Uirf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667

```

RESULT 8
US-09-535-521-24
; Sequence 24, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 24
; LENGTH: 369
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (369)
US-09-535-521-24

```

	Query Match	8.4%;	Score 62.8;	DB 4;	Length 369;
	Best Local Similarity	57.7%;	Pred. No. 3.7e-10;		
	Matches 112;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;
QY	525	GAATCTCATCAGGAGGAGCCTTCTCTGGGCAATCACTGATGAGAAGACAGAGGGCAGTT	584		
Ddb	165	GTATGCCAACAGAAAGGGGCACCTGTGATTGGCCCTCCGGGACCTTGGACAGAGAGGGGAGTT	224		
QY	585	TGTGGATCTGCAGGAAATAGACTGACCTTACACAAACTTGAACGAGGGTGAACCCCAACAA	644		
Ddb	225	TATCTGGATGACGAGAACCCCTTGAACCTATATAGCAACTCGCGCCCGGGGAGGCCCAACAA	284		

Qy	645	TGCTGGTCTTGATCAAGATTGTGTAATGCTACTGAAAAATGGCCAGTGGCAATGACGTCCC	704
Dβ	285	CGGGGGCCAGGGCGAGGACTCGCTGATGATCGAGGCTCGGGGCACTGGAAATGACGCCCTT	344

QY 705 CTGCTCCACCTCCC 718
db 345 CTGCGGCAGCTCGC 358

```

RESULT 9
US-09-535-521-26/c
; Sequence 26, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; DATE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF

```

```

; FILE REFERENCE: AD-3
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 369
; TYPE: DNA
; ORGANISM: Canis familiaris
; US-09-535-521-26

```

Query Match	8.4%;	Score 62.8;	DB 4;	Length 369;
Best Local Similarity	57.7%;	Pred. No. 3.7e-10;		
Matches 112;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;
Qy	525	GAATCTCATCAAGGAGGAAGCCTTCCTGGGCATCACTGATGAGAAGACAGAAAGGGCAGTT	584	
Db	205	GTATGCCAACAAAGAGGGCCTTGGATTGGCTCCGGACCTGGACAGAGAGGGGGAGTT	146	
Qy	585	TGTGGATCTGACAGAAATAGACTGACCTACACAACTGGAAACGAGGGTGAAACCCACAA	644	
Db	145	TATCTGGATGGACGAGAACCCCTTGAACTATAGCAACTGGCGGCCCGGGAGCGCCACAA	86	
Qy	645	TGCTGGTTCATGAAGATTCTGTATTGTCTACTGAAAATGCCAGGTGGGATGACGTCCC	704	
Db	85	CGGGGGCCAGGGCAGGAGCTCCGTGATGATGACAGGGTCGGGGCAGTGGAAATGACGCCTT	26	
Qy	705	CTGCTCCACCTCCC	718	
Db	25	CTCGGCAGCTCGC	12	

```

; NAME/KEY: CDS
; LOCATION: {1}..(384)
US-09-535-521-7

```

	Query Match	8.4%;	Score 62.8;	DB 4;	Length 384;
	Best Local Similarity	57.7%;	Pred. No. 3.8e-10;		
	Matches 112;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;
Qy	525	GAATCTCATCAAGGAGGAAGCCTTCCTGGGCATCACTGATGAGAAGACAGAAGGGCAGTT	584		
Db	156	GTATGCCAACAGAAGGGCACCTGGATTGGCTCCGGGACCTGCACAGAGAGGGGGAGTT	215		
Qy	585	TGTGGATCTCAGCAGGAATAGACTGACCTACACAAACTGGAACGAGGGTGAACCCACAA	644		
Db	216	TATCTGGATGGATCAGAAACCCCTGAACTATAGCACTGGCGGCCGGGGAGGCCAACAA	275		
Qy	645	TGCTGGTTCTGTGTAAGATTGTGTATTGCTACTGAAAAATGGCCAGTGGGAATGACGTCCC	704		
Db	276	CGGGGGCCGCGGAGGACTCGTGTATGATGCAGGGCTCGGGCAGTGGGAATGACGCCTT	335		
Qy	705	CTGCTCCACCTCCC	718		
Db	336	CTCGGCAGCTCGC	349		

```

RESULT 11
US-09-535-521-9/c
; Sequence 9, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-535-521-9

```

	Query Match	8.4%;	Score 62.8;	DB 4;	Length 384;
	Best Local Similarity	57.7%;	Pred. No. 3.8e-10;		
	Matches 112;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;
QY	525	GAATCTCATCAAGGAGGAAGCCTTCCTGGGGCATCCTGATGAGAAGACAGAAAGGGCGAGTT	584		
DB	229	GTATGCCAACAGAAAGGCACCTTGGATTGGCTCCGGGACCTGGACAGAGAGGGGGAGTT	170		
QY	585	TGTGGATCTGACAGGAAATAGACTGACCTTACCAAACTGGAAACGAGGGTGAACCCACAA	644		
DB	169	TATCTGGATGGACAGAGAACCCCTTGAATATAGCAACTGGCGGCCCGGGGAGGCCACAA	110		
QY	645	TGCTGGTTCTGATGAAGATTGTGTATTGCTACTGAAAAAATGGCCAGGTGGAAATGACGTCCC	704		
DB	109	CGGGGGCCAGGGCGGAGGACTGCTGATGATGACAGGGCTCGGGGCACTGGAAATGACGCCTT	50		
QY	705	CTGCTCCACCTCCC	718		
DB	49	CTCGGGCAGCTCGC	36		

RESULT 12
US-09-535-521-10
; Sequence 10, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:

; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; EARLIER FILING DATE: 2000-03-24
; EARLIER FILING DATE: 2000-03-24
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(417)
US-09-535-521-10

Query Match 8.4%; Score 62.8; DB 4; Length 417;
Best Local Similarity 57.7%; Pred. No. 4e-10;
Matches 112; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 525 GAATCTCATCAAGAGGAGCCTTCTGGGCATCACTGATGAGAGACAGAGGCGAGTT 584
Db 189 GTATGCCAACAAAGAGGACCTTGGATTGGCTCCGGGACCTGGACAGAGAGGGGAGTT 248
QY 585 TGTGGATCTGACAGGAATAGACTGACCTACACAACTGGAACAGAGGTTGAAACCAACAA 644
Db 249 TATCTGGATGGACGAGAACCCCTTGAATAGCACTATAGCACTGGCGGCCGGGAGCCCAACAA 308
QY 645 TGCTGGTCTTGATGAAGATTGTGATTGCTACTGAAAAATGGCCAGTGGATGACGTCCC 704
Db 309 CGGGGGCCAGGCGAGGACTGCTGATGATGACAGGGCTCGGGCGAGTGGATGACGCTT 368
QY 705 CTGCTCCACCTCCC 718
Db 369 CTGGCGAGCTCGC 382

RESULT 13
US-09-535-521-12/c
; Sequence 12, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; EARLIER FILING DATE: 2000-03-24
; EARLIER FILING DATE: 2000-03-24
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-535-521-12

Query Match 8.4%; Score 62.8; DB 4; Length 417;
Best Local Similarity 57.7%; Pred. No. 4e-10;
Matches 112; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 525 GAATCTCATCAAGAGGAGCCTTCTGGGCATCACTGATGAGAGACAGAGGCGAGTT 584
Db 229 GTATGCCAACAAAGAGGACCTTGGATTGGCTCCGGGACCTGGACAGAGAGGGGAGTT 170
QY 585 TGTGGATCTGACAGGAATAGACTGACCTACACAACTGGAACAGAGGTTGAAACCAACAA 644

Db 169 TATCTGGATGGACGAGAACCCCTGAACTATAGCAACTGGCGGCCGGGAGCCCAACAA 110
QY 645 TGCTGGTCTTGATGAAGATTGTGATTGCTACTGAAAAATGGCCAGTGGATGACGTCCC 704
Db 109 CGGGGGCCAGGCGAGGACTGCGTGTGATGATGACAGGGCTCGGGGAGTGGATGACGCTT 50
QY 705 CTGCTCCACCTCCC 718
Db 49 CTGGCGAGCTCGC 36

RESULT 14
US-09-535-521-13
; Sequence 13, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; EARLIER FILING DATE: 2000-03-24
; EARLIER FILING DATE: 2000-03-24
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(423)
US-09-535-521-13

Query Match 8.4%; Score 62.8; DB 4; Length 423;
Best Local Similarity 57.7%; Pred. No. 4e-10;
Matches 112; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 525 GAATCTCATCAAGAGGAGCCTTCTGGGCATCACTGATGAGAGACAGAGGCGAGTT 584
Db 195 GTATGCCAACAAAGAGGACCTTGGATTGGCTCCGGGACCTGGACAGAGAGGGGAGTT 254
QY 585 TGTGGATCTGACAGGAATAGACTGACCTACACAACTGGAACAGAGGTTGAAACCAACAA 644
Db 255 TATCTGGATGGACGAGAACCCCTGAACTATAGCACTGGCGGCCGGGAGCCCAACAA 314
QY 645 TGCTGGTCTTGATGAAGATTGTGATTGCTACTGAAAAATGGCCAGTGGATGACGTCCC 704
Db 315 CGGGGGCCAGGCGAGGACTGCGTGTGATGATGACAGGGCTCGGGGAGTGGATGACGCTT 374
QY 705 CTGCTCCACCTCCC 718
Db 375 CTGGCGAGCTCGC 388

RESULT 15
US-09-535-521-15/c
; Sequence 15, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; EARLIER FILING DATE: 2000-03-24
; EARLIER FILING DATE: 2000-03-24
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 15
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Ca
US-09-535-521-15

```

	Query Match	8.4%	Score 62.8;	DB 4;	Length 423;
	Best Local Similarity	57.7%;	Pred. No. 4e-10;		
	Matches 112;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;
Qy	525	GAATCTCATCAAGGAGGAAGCCTTCCTGGGCATCACCTGATGAGAAGACAGAAGGGCGAGTT	584		
Db	229	GTATGCCAACAGAAAGGCACCTGGATTGSCCTCCGGGACCTGGACAGAGAGGGGGAGTT	170		
Qy	585	TGTGGATCTGACAGGAAATGACTGACCTACACAACTGGAAAGAGGGTGAACCCACAA	644		
Db	169	TATCTGGATGGACAGAGAACCCTCGAATATAGCAACTGGCGGCCCGGGAGGCCACAA	110		
Qy	645	TGCTGGTTCTGATCAAGATTGTGTATTGCTACTGAAAAATGGCCAGTGGGAATGACGTCCC	704		
Db	109	CGGGGGCCAGGGCGAGACTCGTGATGTCAGGGCTCGGGGCAGTGGATGACGCCTT	50		
Qy	705	CTGCTCCACCTCCC	718		
Db	49	CTGGCGAGCTCGC	36		

Search completed: June 21, 2004, 19:07:32
Job time : 82 secs

B/GK

QY 535 AAGGAGGAGGCTTCTCTGGGATCACTGATGAGAGAGACAGAGGCGAGTTTGTGATCTG 594
DB 553 ACTGAGAGGCTTCTCTGGGATCACTGATGAGAGAGCTGAGGCGAATTTGTGATCTG 612
QY 595 ACAGGAAATAGACTGACCTTACAAACTGGAACGAGGGTGAACCCCAACAATGCTGGTCT 654
DB 613 ACAGGAAAGGGGTTGACCTTACCAAACTGGAATGATGGCGAGCTTAAACAACGTTCTCT 672
QY 655 GATGAGATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 714
DB 673 GGGGAGCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 732
QY 715 TCCCATCTGGCGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 747
DB 733 TCCTTTGTGACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 765

RESULT 7

US-09-917-800A-1710
; Sequence 1710, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1710
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_022704
US-09-917-800A-1710

Query Match 48.8%; Score 364.6; DB 9; Length 1037;
Best Local Similarity 70.7%; Pred. No. 1.1e-109;
Matches 483; Conservative 0; Mismatches 200; Indels 0; Gaps 0;

QY 65 CTGTGACCTGTGTGAGGATGCCAAAGACCTGCCCTGACGTGATGCTGTAGCTTCCAG 124
DB 233 CGGAGACCTTAAACCGAAGGGCTCAAAGTAGTGCCTGTGATTGCTTGCAGTTCTCCGG 292
QY 125 GCATCAACGGCTTCCAGCAAGATGGGGTGATGCGCCACCAAGGAGAGAAAGGGGGAAC 184
DB 293 CTGTGAACGGCTTCCAGGCAAGATGGACACACGCTGTCCAAAGGGAGAGAAAGGGGAAC 352
QY 185 CAGGCCAAGGGCTCAGAGGCTTACAGGGCCCCCTTGAAAGTTGGGGCCCTCCAGGAATC 244

DB 353 CGGTTCAAGGCTCAGAGGCTTGCAGGCGCTCTCTGTGAAAAGTAGGACCTGCAGGCGCCC 412
QY 245 CAGGCGCTTCTGGGTCAACAGGACCAAAAGGGCCAAAAGGAGACCTGGAAAAGTCCGG 304
DB 413 CAGGGAATCCTGGGTCAAAGGAGCAACGGGACCAAAAGGAGACCGTGGAGAGAGTGTAG 472
QY 305 ATGGTGATAGTAGCTGCTCCTCAGAAAAGAAAAGCTCTGCAAAACAGAAAATGGCACGTA 364
DB 473 AATTGATATCTACCAACATTTGATTTAGAAATTTGAGCCCTGCGATCGGAGCTGAGAGCTA 532
QY 365 TCAAAAGTGGCTGACCTTCTCTCTGGGCAACAAAGTTGGGAAACAAGTTCTTCTTGACCA 424
DB 533 TGAGAAAGTGGGTGCTCTCTCTATGAGTGAAAATGTTGGAAAAGAGTACTTTCATGAGCA 592
QY 425 ATGGTGAATATGATGACCTTTGAAAAGTGAAGGCGCTTGTGTCTCAAGTTCCAGGCTCTG 484
DB 593 GTGTTAGAAGGATGCCCTTAAACAGAGCGAAGGCTCTGTGTCTCCGAACTCCAGGGCACTG 652
QY 485 TGGCCACCCCGAGGAATGCTGCAGAGAATGGAGCCATTCAAGATCTCATCAAGGAGGAAG 544
DB 653 TGGCCACTCCCGAGGAATGCTGAGGAAAATAGGGCCATTCAGAAATGTGSCCAAGATTTG 712
QY 545 CTTCTCTGGGCATCACTCATGAGAAAGACAGAGGCGCAGTTTGTGTGATCTGACAGAAATA 604
DB 713 CTTCTCTGGGCATAACGAGCAGAGGACTGAAAACGTTTTTGGAGGACCTGACAGGAAACA 772
QY 605 GACTGACCTACAAACTGGAAGAGGGTGAACCCCAACAATGCTGTGTTCTGATGAAT 664
DB 773 GAGTGGCTTACACTAACTGGAATGAGGGTGAAGCCCAACAATGTGGGCTCTGGGAAAAC 832
QY 665 GTGATTGCTACTGAAAATGSCCAGTGAATGAGTCCCTGCTCCCACTCCCATCTCG 724
DB 833 GTGTGTGCTCTTGACAAATGGGAAGTGAATGAGTCTCTGCTGATTCCTTTTGG 892
QY 725 CCGTCTGTGAGTCCCTATCTGA 747
DB 893 TAGTTGTGAAATCTCTGACTGA 915

RESULT 8

US-10-388-934-107
; Sequence 107, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 107
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-388-934-107

Query Match 38.5%; Score 287.6; DB 16; Length 1068;
Best Local Similarity 68.1%; Pred. No. 3.6e-84;
Matches 417; Conservative 0; Mismatches 189; Indels 6; Gaps 1;

QY 142 GSCAAAGATGGCGGTGATGGCACCAAGGAGAAAAGGGGGAACCGGCAAGGGCTCAGA 201
DB 64 GGCAGAGACGGGAGAGATGGGCCCAAGGGGAGAGGAGAACCAAGGTCAAGGGCTCAG 123
QY 202 GCGTTACAGGGCCCCCTTGAAAAGTTGGGGCCCTCCAGGAAATCCAGGGGCTTCTGGGTCA 261

QY 1 ATGTCCTGTTTCCATCACTCCCTCTCTCTCTGAGTATGTCGACGCTTACTCA 60
|||||
Db 715 ATGTCCTGTTTCCATCACTCCCTCTCTCTCTGAGTATGTCGACGCTTACTCA 774
|||||
QY 61 GAACTGTGACCTGTGAGGATGCCAAAGACCTGCCCTGCAGTGAATGCTGTAGCTCT 120
|||||
Db 775 GAACTGTGACCTGTGAGGATGCCAAAGACCTGCCCTGCAGTGAATGCTGTAGCTCT 834
|||||
QY 121 CCAGGCATCAACGGCTTCCAGGCAAGATGGCGTGTATGGCAACCAAGGAGAAAGGGG 180
|||||
Db 835 CCAGGCATCAACGGCTTCCAGGCAAGATGGCGTGTATGGCAACCAAGGAGAAAGGGG 894
|||||
QY 181 GAACCAAG 188
|||||
Db 895 GAACCAAG 902
|||||

RESULT 11
US-10-070-415A-49
; Sequence 49, Application US/10070415A
; Publication No. US20040043379A1
; GENERAL INFORMATION:
; APPLICANT: HASHIMOTO, Koji
; APPLICANT: ASHIMOTO, Michie
; APPLICANT: MISHIRO, Shunji
; APPLICANT: OOTA, Yasuhiko
; TITLE OF INVENTION: DETECTION OF NUCLEIC ACID ASSOCIATED WITH DISEASE
; FILE REFERENCE: 220633US2SRDPCT
; CURRENT APPLICATION NUMBER: US/10/070,415A
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/JP02/02030
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: JP 2001-090053
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: JP 2001-284112
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 1802
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (96)..(96)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (425)..(425)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (649)..(649)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (884)..(884)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (875)..(875)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
US-10-070-415A-49

Query Match 24.9%; Score 186; DB 13; Length 1802;
Best Local Similarity 98.9%; Pred. No. 2.1e-50;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGTCCTGTTTCCATCACTCCCTCTCTCTCTGAGTATGTCGACGCTTACTCA 60
|||||
Db 715 ATGTCCTGTTTCCATCACTCCCTCTCTCTCTGAGTATGTCGACGCTTACTCA 774
|||||
QY 61 GAACTGTGACCTGTGAGGATGCCAAAGACCTGCCCTGCAGTGAATGCTGTAGCTCT 120

Db 775 GAACTGTGACCTGTGAGGATGCCAAAGACCTGCCCTGCAGTGAATGCTGTAGCTCT 834
|||||
QY 121 CCAGGCATCAACGGCTTCCAGGCAAGATGGCGTGTATGGCAACCAAGGAGAAAGGGG 180
|||||
Db 835 CCAGGCATCAACGGCTTCCAGGCAAGATGGCGTGTATGGCAACCAAGGAGAAAGGGG 894
|||||
QY 181 GAACCAAG 188
|||||
Db 895 GAACCAAG 902
|||||

RESULT 12
US-10-070-415A-53
; Sequence 53, Application US/10070415A
; Publication No. US20040043379A1
; GENERAL INFORMATION:
; APPLICANT: HASHIMOTO, Koji
; APPLICANT: ASHIMOTO, Michie
; APPLICANT: MISHIRO, Shunji
; APPLICANT: OOTA, Yasuhiko
; TITLE OF INVENTION: DETECTION OF NUCLEIC ACID ASSOCIATED WITH DISEASE
; FILE REFERENCE: 220633US2SRDPCT
; CURRENT APPLICATION NUMBER: US/10/070,415A
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/JP02/02030
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: JP 2001-090053
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: JP 2001-284112
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 1802
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (96)..(96)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (425)..(425)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (649)..(649)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (868)..(868)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (875)..(875)
; OTHER INFORMATION: n is a nucleotide selected from a, g, c, or t
US-10-070-415A-53

Query Match 24.9%; Score 186; DB 13; Length 1802;
Best Local Similarity 98.9%; Pred. No. 2.1e-50;
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGTCCTGTTTCCATCACTCCCTCTCTCTCTGAGTATGTCGACGCTTACTCA 60
|||||
Db 715 ATGTCCTGTTTCCATCACTCCCTCTCTCTCTGAGTATGTCGACGCTTACTCA 774
|||||
QY 61 GAACTGTGACCTGTGAGGATGCCAAAGACCTGCCCTGCAGTGAATGCTGTAGCTCT 120
|||||
Db 775 GAACTGTGACCTGTGAGGATGCCAAAGACCTGCCCTGCAGTGAATGCTGTAGCTCT 834
|||||
QY 121 CCAGGCATCAACGGCTTCCAGGCAAGATGGCGTGTATGGCAACCAAGGAGAAAGGGG 180
|||||

B/GNK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 18, 2004, 20:53:21 ; Search time 15 Seconds
(without alignments)
5186.187 Million cell updates/sec

Title: US-10-054-536-2

Perfect score: 1365
Sequence: 1 atgtccctgtttccatcact.....tctgtgagttccctatctga 747

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DRV=xlp
-Q/cgn2.1/USPTO.spool.p/US10054536/runat.18062004.173455.13573/app_query.fasta.1.903
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=resp -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10054536 @CGN.1.1.16 @runat.18062004.173455.13573 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1321	96.8	248	1 MABC_HUMAN	P11226 homo sapien
2	847.5	62.1	249	1 MABC_BOVIN	O02659 bos taurus
3	790	57.9	244	1 MABC_MOUSE	P41317 mus musculus
4	785	57.5	244	1 MABC_RAT	P08661 rattus norv
5	711.5	52.1	239	1 MABA_MOUSE	P39039 mus musculus
6	705.5	51.7	238	1 MABA_RAT	P19999 rattus norv
7	454	33.3	375	1 PSPD_HUMAN	P35247 homo sapien
8	449.5	32.9	371	1 CL46_BOVIN	Q8mbz9 bos taurus
9	441.5	32.3	371	1 CONG_BOVIN	P23805 bos taurus
10	441	32.3	374	1 PSPD_RAT	P50448 rattus norv
11	439	32.2	374	1 PSPD_MOUSE	P52404 mus musculus
12	427	31.3	321	1 CL43_BOVIN	P42916 bos taurus
13	417.5	30.6	247	1 PSPA_RABIT	P12842 corytolagus
14	415.5	30.4	369	1 PSPD_BOVIN	P35246 bos taurus
15	405.5	29.7	247	1 PSPA_CAVPO	P50403 cavia porce
16	404	29.6	248	1 PSPA_RAT	P08427 rattus norv
17	393.5	28.8	248	1 PSPA_HUMAN	P07714 homo sapien
18	393	28.8	248	1 PSPA_MOUSE	P35242 mus musculus

19	393	28.8	249	1	PSPA_PIG	P49874 sus scrofa
20	374.5	27.4	248	1	PSPA_CANFA	P06908 canis famil
21	204	14.9	197	1	CLE1_HUMAN	O75596 homo sapien
22	201.5	14.8	331	1	FCB2_MOUSE	P20593 mus musculus
23	199.5	14.6	207	1	LECH_CHICK	P02707 gallus gall
24	193.5	14.2	253	1	C1QB_MOUSE	P14106 mus musculus
25	192.5	14.1	253	1	C1QB_RAT	P31721 rattus norv
26	192.5	14.1	335	1	FCN1_RAT	Q9wtc8 rattus norv
27	192	14.1	244	1	APM1_HUMAN	Q15848 homo sapien
28	189.5	13.9	247	1	APM1_MOUSE	Q60994 mus musculus
29	189	13.8	245	1	C1QA_MOUSE	P98086 mus musculus
30	188	13.8	1670	1	CA34_HUMAN	O01955 homo sapien
31	187	13.7	334	1	FCN1_MOUSE	O70165 mus musculus
32	187	13.7	1763	1	CA24_ASCSU	P27393 ascaris suu
33	187	13.7	1888	1	CA1E_CHICK	P32018 gallus gall
34	186	13.6	1890	1	CA44_HUMAN	P53420 homo sapien
35	185.5	13.6	202	1	TEIN_HUMAN	P05452 homo sapien
36	185	13.6	912	1	PGCB_BOVIN	Q28062 bos taurus
37	184	13.5	251	1	C1QB_HUMAN	P02746 homo sapien
38	183.5	13.4	245	1	C1QA_HUMAN	P02745 homo sapien
39	183	13.4	675	1	CA39_CHICK	P32017 gallus gall
40	183	13.4	688	1	CA29_MOUSE	O07543 mus musculus
41	182.5	13.4	245	1	C1QC_HUMAN	P02747 homo sapien
42	182.5	13.4	246	1	C1QC_MOUSE	Q02105 mus musculus
43	181.5	13.3	1329	1	KF10_HUMAN	Q9p218 homo sapien
44	181.5	13.3	1459	1	CA12_MOUSE	P28481 mus musculus
45	181	13.3	689	1	CA29_HUMAN	Q14055 homo sapien

ALIGNMENTS

RESULT 1

ID	MABC_HUMAN	STANDARD	PRT	248 AA
AC	P11226; O86S14; Q96K84; Q96TF7; Q96TF8; Q96TF9;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Mannose-binding protein C precursor (MBP-C) (MBP1) (Mannan-binding protein) (Mannose-binding lectin).			
DE	MBL2 OR MBL.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=90010778; PubMed=2477486;			
RA	Sastry K., Herman G.A., Day L.E., Deignan E., Bruns G., Morton C.C.,			
RA	Ezekowitz R.A.B.;			
RT	"The human mannose-binding protein gene. Exon structure reveals its evolutionary relationship to a human pulmonary surfactant gene and localization to chromosome 10.";			
RL	J. Exp. Med. 170:1175-1189(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88171281; PubMed=2450948;			
RA	Ezekowitz R.A.B., Day L.E., Herman G.A.;			
RT	"A human mannose-binding protein is an acute-phase reactant that shares sequence homology with other vertebrate lectins.";			
RL	J. Exp. Med. 167:1034-1046(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90073571; PubMed=2590164;			
RA	Taylor M.E., Brickell P.M., Craig R.K., Summerfield J.A.;			
RT	"Structure and evolutionary origin of the gene encoding a human serum mannose-binding protein.";			
RL	Biochem. J. 262:763-771(1989).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANTS CYS-52; ASP-54 AND GLU-57.			
RX	MEDLINE=98414317; PubMed=9743385;			
RA	Madsen H.O., Satz M.L., Høgh B., Svejgaard A., Garred P.;			

DR EMBL; U09015; AAA2010.1; JOINED.
 DR EMBL; D11440; BAA02005.1; -.
 DR EMBL; BC010760; AAH0760.1; -.
 DR PIR; I48651; LNM5MC.
 DR HSSP; P08661; IRDO.
 DR MGD; MGI:96924; Mb12.
 DR InterPro; IPR008160; Collagen.
 DR InterPro; IPR001304; Lectin_C.
 DR Pfam; PF001391; Collagen; 1.
 DR Pfam; PF00059; lectin_c; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
 DR PROSITE; PS00041; C-TYPE LECTIN 2; 1.
 KW Complement pathway; Membrane; Mannose-binding; Calcium; Repeat;
 FT SIGNAL 1
 FT CHAIN 18
 FT DOMAIN 19 244
 FT DOMAIN 38 96
 FT DOMAIN 129 241
 FT MOD_RES 43 43
 FT MOD_RES 58 58
 FT MOD_RES 69 69
 FT MOD_RES 78 78
 FT MOD_RES 81 81
 FT DISULFID 29 29
 FT DISULFID 34 34
 FT DISULFID 151 240
 FT DISULFID 218 232
 FT CONFLICT 3 3
 FT CONFLICT 15 15
 FT CONFLICT 15 15
 SQ SEQUENCE 244 AA; 25957 MW; 49AE84E2290DEB0A CRC64;

Alignment Scores:
 Pred. No.: 3.63e-58 Length: 244
 Score: 790.00 Matches: 150
 Percent Similarity: 75.90% Conservative: 39
 Best Local Similarity: 60.24% Mismatches: 50
 Query Match: 57.88% Indels: 10
 DB: 1 Gaps: 5

US-10-054-536-2 (1-747) x MABC_MOUSE (1-244)

QY 1 ATGTCCTGTTTCATCATCTCCCTCTCTCTCTGATGATGGCAGGCTTACTCA 60
 Db 1 MetSerIlePheThrSer-----PheLeuLeuLeuLeuCysValValThrValValTyrAla 18
 QY 61 GAACTGTGACCTGTGAGGATGCCAAAGACCTGCGCTGACGTGATGCTGAGCTCT 120
 Db 19 GluThrLeuThr---GluGlyValGlnAsnSerCysPro---ValValThrCysSerSer 36
 QY 121 CCAGGCATCAACGGCTCCAGCAAGATGGCGTGGTGGTGGTGGTGGTGGTGGTGG 180
 Db 37 ProGlyLeuAsnGlyPheProGlyLysAspGlyArgAspGlyAlaLysGlyGlyLysGly 56
 QY 181 GAACAGGCGCAAGGGCTCAGAGCTTACAGGGCCCCCTCGAAAGTTG-----GGG 231
 Db 57 GluProGlyGlnGlyLeuArgGlyLeuGlnGlyProProGlyLysValGlyProThrGly 76
 QY 232 CTTCCAGGAATCAGGGCCTTCTGGTCCACAGGACCAAGGCCCAAAAGGAGGACCCCT 291
 Db 77 ProProGlyAsnProGlyLeuLysGlyAlaValGlyProLysGlyAspArgGlyAsp--- 95
 QY 292 GGAAGAGTCCGATGGTGTAGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
 Db 96 -----ArgAlaGluPheAspThrSerGluIleAspSerGluIleAlaLeuArgSer 113
 QY 352 GAAATGACGCTATCAAAAAGTGGCTGACCTCTCTCTGGCGCAACAAAGTTGGCAACAG 411
 Db 114 GluLeuArgAlaLeuArgAsnTrpValLeuPheSerLeuSerGluLysValGlyLysLys 133
 QY 412 TTTCTCTGACCAATGGTGAATATGACCTTTTCAAAAAGTGAAGCCCTTGTGTGCAAG 471
 Db 134 TyrPheValSerSerValLysLysMetSerLeuAspArgValLysAlaLeuLysCysSerGlu 153

QY 472 TTCAGGCCTCTGTGGCCACCCAGGAATGCTGCAGAGAAATGGAGGCATTCAGAAATCTC 531
 Db 154 PheGlnGlySerValAlaThrProArgAsnAlaGluGluAsnSerAlaIleGlnLysVal 173
 QY 532 ATCAAGGAGGAAGCCTTCCTGGGCATCCTGATGAGAGACAGAAAGGCAGTGTGTGGAT 591
 Db 174 AlalysAspIleAlaTyrLeuGlyIleThrAspValArgValGluGlySerPheGluAsp 193
 QY 592 CTGACAGGAATAGACTGACCTACACAACTGGAACGAGGGTGAACCAACAAATGCTGGT 651
 Db 194 LeuThrGlyAsnArgValArgTyrThrAsnTrpAsnAspGlyGluProAsnAsnThrGly 213
 QY 652 TCTGATGACAGATGCTGTATTGCTACTGAAAAATGGCCAGTGGATGACGTCCTGCTCC 711
 Db 214 AspGlyGluAspCysValValIleLeuGlyAsnGlyLysTrpAsnAspValProCysSer 233
 QY 712 ACCTCCCATCTGGCGCTCTGTGAGTTC 738
 Db 234 AspSerPheLeuAlaIleCysGluPhe 242

RESULT 4
 MABC RAT STANDARD; PRT; 244 AA.
 ID MABC RAT
 AC P08661;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Mannose-binding protein C precursor (MBP-C) (Mannan-binding protein)
 DE (RA-reactive factor P28A subunit) (RARP/P28A).
 GN MBL2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92299655; PubMed=1607365;
 RA Wada M., Itoh N., Ohta M., Kawasaki T.;
 RT "Characterization of rat liver mannan-binding protein gene.";
 RL J. Biochem. 111:66-73(1992).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=86196130; PubMed=3009480;
 RA Drickamer K., Dordal M.S., Reynolds L.;
 RT "Mannose-binding proteins isolated from rat liver contain
 RT carbohydrate-recognition domains linked to collagenous tails. Complete
 RT primary structures and homology with pulmonary surfactant
 RT apoprotein.";
 RL J. Biol. Chem. 261:6878-6887(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=87194686; PubMed=3032924;
 RA Oka S., Itoh N., Kawasaki T., Yamashina I.;
 RT "Primary structure of rat liver mannan-binding protein deduced from
 RT its cDNA sequence.";
 RL J. Biochem. 101:135-144(1987).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 133-244.
 RX MEDLINE=96132792; PubMed=8557671;
 RA Ng K.K.-S., Drickamer K., Weis W.I.;
 RT "Structural analysis of monosaccharide recognition by rat liver
 RT mannose-binding protein.";
 RL J. Biol. Chem. 271:663-674(1996).
 CC -!- FUNCTION: Binds mannose and N-acetylglucosamine in a calcium-
 CC dependent manner. Is capable of host defense against pathogens, by
 CC activating the classical complement pathway independently of the
 CC antibody.
 CC -!- SUBUNIT: Oligomeric complex of 6 set of homotrimers.
 CC -!- SUBCELLULAR LOCATION: THIS PROTEIN IS LOCATED INTRACELLULAR, MOST

IO MABA RAT STANDARD; PRT; 238 AA.
AC P19999;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Mannose-binding protein A precursor (MBP-A) (Mannan-binding protein).
DN MBP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=86196130; PubMed=3009480;
RA Drickamer K., Dordal M.S., Reynolds L.;
RT "Mannose-binding proteins isolated from rat liver contain
RT carbohydrate-recognition domains linked to collagenous tails.
RT Complete primary structures and homology with pulmonary surfactant
RT apoprotein.";
RL J. Biol. Chem. 261:6878-6887(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=87137502; PubMed=3029088;
RA Drickamer K., McCreary V.;
RT "Exon structure of a mannose-binding protein gene reflects its
RT evolutionary relationship to the asialoglycoprotein receptor and
RT nonfibrillar collagens.";
RL J. Biol. Chem. 262:2582-2589(1987).
RN [3]
RP SEQUENCE OF 18-42.
RX MEDLINE=87222358; PubMed=3584121;
RA Ikeda K., Sannoh T., Kawasaki N., Kawasaki T., Yamashina I.;
RT "Serum lectin with known structure activates complement through the
RT classical pathway.";
RL J. Biol. Chem. 262:7451-7454(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 124-238.
RX MEDLINE=92086855; PubMed=1721241;
RA Weis W.I., Kahn R., Fourme R., Drickamer K., Hendrickson W.A.;
RT "Structure of the calcium-dependent lectin domain from a rat mannose-
RT binding protein determined by MAD phasing.";
RL Science 254:1608-1615(1991).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 124-238.
RX MEDLINE=93063338; PubMed=1436090;
RA Weis W.I., Drickamer K., Hendrickson W.A.;
RT "Structure of a C-type mannose-binding protein complexed with an
RT oligosaccharide.";
RL Nature 360:127-134(1992).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 84-238.
RX MEDLINE=95219384; PubMed=7704532;
RA Weis W.I., Drickamer K.;
RT "Trimeric structure of a C-type mannose-binding protein.";
RL Structure 2:1227-1240(1994).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 90-238.
RX MEDLINE=99119227; PubMed=9922165;
RA Ng K.K.-S., Park-Snyder S., Weis W.I.;
RT "Ca2+-dependent structural changes in C-type mannose-binding
RT proteins.";
RL Biochemistry 37:17965-17976(1998).
CC -1- FUNCTION: Binds mannose and N-acetylglucosamine in a calcium-
CC dependent manner. Is capable of host defense against pathogens, by
CC activating the classical complement pathway independently of the
CC antibody.
CC -1- SUBUNIT: Oligomeric complex of 6 set of homotrimers.
CC -1- SUBCELLULAR LOCATION: THIS PROTEIN IS LOCATED INTRACELLULAR, MOST
CC PREDOMINANTLY IN THE ROUGH ENDOPLASMIC RETICULUM AND IN THE GOLGI
CC APPARATUS, AND LITTLE OR NO MBP IS PRESENT IN PLASMA MEMBRANES AND
CC LYSOSOMES. THE MBP IS EXCLUSIVELY LOCALIZED IN THE CISTERNA SPACE

CC OF THE ORGANELLES, PROBABLY AS A LOOSELY BOUND MEMBRANE PROTEIN.
CC -1- SIMILARITY: Contains 1 collagenous domain.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M14105; AAA98781.1; --
DR EMBL; M14104; AAA98781.1; JOINED.
DR PIR; B24791; LNRMTA.
DR PDB; 1MSB; 15-JAN-92.
DR PDB; 2MSB; 31-OCT-93.
DR PDB; 1AFA; 03-APR-96.
DR PDB; 1AFB; 03-APR-96.
DR PDB; 1AFD; 03-APR-96.
DR PDB; 1RTM; 07-FEB-95.
DR PDB; 1KMB; 12-FEB-97.
DR PDB; 2KMB; 12-FEB-97.
DR PDB; 3KMB; 12-FEB-97.
DR PDB; 4KMB; 12-FEB-97.
DR PDB; 1YTT; 10-JUN-96.
DR PDB; 1BCH; 17-JUN-98.
DR PDB; 1BCU; 17-JUN-98.
DR PDB; 1BUU; 09-SEP-98.
DR PDB; 1FIF; 10-JAN-01.
DR PDB; 1FIH; 10-JAN-01.
DR PDB; 1KWT; 05-JUL-02.
DR PDB; 1KWU; 05-JUL-02.
DR PDB; 1KWV; 05-JUL-02.
DR PDB; 1KWW; 05-JUL-02.
DR PDB; 1KWY; 05-JUL-02.
DR PDB; 1KWZ; 05-JUL-02.
DR PDB; 1KX0; 05-JUL-02.
DR PDB; 1KX1; 05-JUL-02.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01391; Collagen; 1.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLSCT_1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
KW Complement pathway; Membrane; Mannose-binding; Calcium; Repeat;
KW Signal; Collagen; Lectin; Glycoprotein; Hydroxylation; 3D-structure.
FT SIGNAL 1 17
FT CHAIN 18 238 MANNOSE-BINDING PROTEIN A.
FT DOMAIN 39 88 COLLAGEN-LIKE.
FT DOMAIN 143 238 C-TYPE LECTIN (SHORT FORM).
FT MOD_RES 43 43 HYDROXYLATION (POTENTIAL).
FT MOD_RES 61 61 HYDROXYLATION.
FT MOD_RES 67 67 HYDROXYLATION.
FT MOD_RES 73 73 HYDROXYLATION.
FT MOD_RES 78 78 HYDROXYLATION (POTENTIAL).
FT DISULFID 145 234
FT DISULFID 212 226
FT CONFLICT 156 156
FT HELIX 91 119
FT TURN 120 121
FT TURN 124 125
FT STRAND 128 136
FT HELIX 138 147
FT TURN 148 149
FT STRAND 151 152
FT HELIX 158 168
FT STRAND 172 177
FT TURN 182 183
FT STRAND 186 187
FT TURN 188 189
R -> K (IN REF. 2).

```

FT STRAND 192 192
FT STRAND 198 198
FT TURN 200 201
FT TURN 207 208
FT STRAND 212 215
FT TURN 217 218
FT STRAND 221 224
FT TURN 226 227
FT STRAND 230 236
SQ SEQUENCE 238 AA; 25308 MW; 1A927482B8A0CB3D CRC64;

Alignment Scores:
Pred. No.: 3,98e-51 Length: 238
Score: 705.50 Matches: 143
Percent Similarity: 70.08% Conservative: 28
Best local Similarity: 58.61% Mismatches: 60
Query Match: 51.68% Indels: 13
DB: 1 Gaps: 4

US-10-054-536-2 (1-747) x MABA_RAT (1-238)
QY 19 CTCCTCTCCTT---CTCTGATGATGGTGGCAGCGCTTACTCAGAACTGTGACCTGT 75
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4 LeuProLeuValLeuLeuValValSerValSerSerSerGlySerGlnThrCys 23
QY 76 GAGGATGCCAAAGACCTGCCCTCAGTGATGCTGTGCTGTAGCTCTCCAGGCATCAACGGC 135
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
24 GluGluThrLeuLeuThrCys---SerValIleAlaCys----- 35
QY 136 TTCCCAGGCAAGATGGCGTGCATGCCAACAGGAGGAGAAAGGGGAAACAGGCCAAGGG 195
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
36 -----GlyArgAspGlyArgAspGlyProLysGlyGluLysGlyGluProGlyGlnGly 53
QY 196 CTCAGAGGCTTACAGGCCCCCTCGAAGTTGGGGCTCCAGAGAAATCCAGGCGCTTCT 255
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
54 LeuArgGlyLeuGlnGlyProGlyProGlyLysLeuGlyProGlySerValGlyAlaPro 73
QY 256 GGGTCACAGGACCAAGGCGCAAGAGGAGACCCCTGGAAAGTCCGATGGTATAGT 315
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 GlySerGlnGlyProLysGlyGlnLysGlyAspArgGlyAspSerArgAlaIleGluVal 93
QY 316 AGCCTGGCT-----GCCTCAGAAAGAAAGCTCTGCAACAGAAATGCCAGTATCAAA 369
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 LysLeuAlaAsnMetGluAlaGluIleAsnThrLeuLysSerLysLeuGluLeuThrAsn 113
QY 370 AAGTGGCTGACCTTCTCTGGGCAACAAAGTTGGGAAACAGTCTCTCTGACCAATGCT 429
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 LysLeuHisAlaPheSerMetGlyLysLysSerGlyLysLysPheValThrAsnHis 133
QY 430 GAAATAAGTACCTTTGAAAGAGTGGAGGCTTGTGTCAAGTTCAGGCTCTGTGGCC 489
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 GluArgMetProPheSerLysValLysAlaLeuLysSerGluLeuArgGlyThrValAla 153
QY 490 ACCCCAGGAATGCTGCAGAGAAATGAGGCATTCAGAAATCTCATCAAGGAGGAGGCTTC 549
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
154 IleProArgAsnAlaGluGluAsnLysAlaIleGlnGluValAlaLysThrSerAlaPhe 173
QY 550 CTGGGATCCTGATGAGAAGACAGAGGCGAGTTGTGGATCTGACAGGAATAGACTG 609
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
174 LeuGlyIleThrAspGluValThrGluGlyGlnPheMetTyrValThrGlyGlyArgLeu 193
QY 610 ACTACACAACTGGACAGGCTGACCAACAAATGCTGCTGCTGATGAGAGTTGTGTA 669
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
194 ThrTyrSerAsnTrpLysAspGluProAsnAspHisGlySerGlyGluAspCysVal 213
QY 670 TTGCTACTGAAATAGCCAGTGGATAGCTGCCCTCTCCACCTCCCATCTGGCCGCTC 729
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 ThrIleValAsnGlnGlyLeuTrpAsnAspIleSerCysGlnAlaSerHisThrAlaVal 233
QY 730 TGTAGTTCCTT 741
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 CysGluPhePro 237

```

```

RESULT 7
PSPD_HUMAN
ID PSPD_HUMAN STANDARD; PRT; 375 AA.
AC P35247;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
GN SFTPD OR SFTPA OR PSPD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93155122; PubMed=8428971;
RA Crouch E., Rust K., Veile R., Donis-Keller H., Grosso L.;
RT "Genomic organization of human surfactant protein D (SP-D). SP-D is
RT encoded on chromosome 10q22.2-23.1.";
RL J. Biol. Chem. 268:2976-2983(1993).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 214-243.
RC TISSUE=Amniotic fluid, and Lung;
RX MEDLINE=92322003; PubMed=1339284;
RA Lu J., Willis A.C., Reid K.B.M.;
RT "Purification, characterization and cDNA cloning of human lung
RT surfactant protein D.";
RL Biochem. J. 284:795-802(1992).
RN [3]
RP SEQUENCE OF 60-375 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=91378578; PubMed=1898081;
RA Rust K., Grosso L., Zhang V., Chang D., Persson A., Longmore W.,
RA Cai G.-Z., Crouch E.;
RT "Human surfactant protein D: SP-D contains a C-type lectin
RT carbohydrate recognition domain.";
RL Arch. Biochem. Biophys. 290:116-126(1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=99197291; PubMed=10368295;
RA Hakansson K., Lim N.K., Hoppe H.-J., Reid K.B.M.;
RT "Crystal structure of the trimeric alpha-helical coiled-coil and the
RT three lectin domains of human lung surfactant protein D.";
RL Structure 7:255-264(1999).
CC -!- FUNCTION: Contributes to the lung's defense against inhaled
CC microorganisms. Binds strongly maltose residues and to a lesser
CC extent other alpha-glucosyl moieties. It could participate in the
CC extracellular reorganization or turnover of pulmonary surfactant.
CC -!- SUBUNIT: Oligomeric complex of 4 set of homotrimers.
CC -!- PTM: The N-terminus is blocked.
CC -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC protein. There are 4 surfactant-associated protein: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC hydrophobic proteins (SP-B and SP-C).
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L05485; AAB59450.1;
CC EMBL; L05483; AAB59450.1; JOINED.
CC EMBL; L05484; AAB59450.1; JOINED.
CC EMBL; X65018; CAA46152.1;
CC PIR; A45225; A45225.
CC PDB; 1B08; 29-NOV-99.

```


RC TISSUE=Liver;
RX MEDLINE=94215917; PubMed=8163202;
RA Liou L.S., Sastry R., Hartshorn K.L., Lee Y.M., Okarma T.B.,
RA Tauber A.I., Sastry K.N.;
RT "Bovine conglutinin (BC) mRNA expressed in liver: cloning and
RT characterization of the BC cDNA reveals strong homology to surfactant
RT protein-D.";
RL Gene 141:277-281(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Semen;
RX MEDLINE=94267222; PubMed=8207234;
RA Liou L.S., Sastry R., Hartshorn K.L., Lee Y.M., Okarma T.B.,
RA Tauber A.I., Sastry K.N.;
RT "Bovine conglutinin gene exon structure reveals its evolutionary
RT relationship to surfactant protein-D.";
RL J. Immunol. 153:173-180(1994).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94128104; PubMed=8297370;
RA Kawasaki N., Itoh N., Kawasaki T.;
RT "Gene organization and 5'-flanking region sequence of conglutinin: a
RT C-type mammalian lectin containing a collagen-like domain.";
RL Biochem. Biophys. Res. Commun. 198:597-604(1994).
RN [6]
RP SEQUENCE OF 21-371.
RX MEDLINE=91131556; PubMed=1993651;
RA Lee Y.-M., Leiby K.R., Allar J., Paris K., Lerch B., Okarma T.B.;
RT "Primary structure of bovine conglutinin, a member of the C-type
RT animal lectin family.";
RL J. Biol. Chem. 266:2715-2723(1991).
RN [7]
RP PRELIMINARY SEQUENCE OF 21-52.
RX MEDLINE=87184551; PubMed=3566740;
RA Young N.M., Leon M.A.;
RT "The carbohydrate specificity of conglutinin and its homology to
RT proteins in the hepatic lectin family.";
RL Biochem. Biophys. Res. Commun. 143:645-651(1987).
CC -!- FUNCTION: Calcium-dependent lectin-like protein which binds to a
CC yeast cell wall extract and immune complexes through the
CC complement component (C3bi). It is capable of binding nonreducing
CC terminal N-acetylglucosamine, mannose, and fucose residues.
CC -!- SUBUNIT: Oligomeric complex of 4 set of homotrimers.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; D14085; BAA03170.1; -;
DR EMBL; X71774; CAA50665.1; -;
DR EMBL; L18871; AAA20126.1; -;
DR EMBL; U08860; AAB60624.1; JOINED.
DR EMBL; U08854; AAB60624.1; JOINED.
DR EMBL; U08855; AAB60624.1; JOINED.
DR EMBL; U08856; AAB60624.1; JOINED.
DR EMBL; U08857; AAB60624.1; JOINED.
DR EMBL; U08858; AAB60624.1; JOINED.
DR EMBL; U08859; AAB60624.1; JOINED.
DR EMBL; D25302; BAA04983.2; JOINED.
DR EMBL; D25296; BAA04983.2; JOINED.
DR EMBL; D25297; BAA04983.2; JOINED.
DR EMBL; D25298; BAA04983.2; JOINED.
DR EMBL; D25299; BAA04983.2; JOINED.
DR EMBL; D25300; BAA04983.2; JOINED.
DR EMBL; D25301; BAA04983.2; JOINED.
DR PIR; I45878;

DR PIR; JN0450; JN0450.
DR HSSP; P35247; 1B08.
DR InterPro; IPR008161; Clq_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01391; Collagen; 3.
DR Pfam; PF00059; Lectin_C; 1.
DR ProDom; PD000007; Clq_helix; 1.
DR SMART; SM00034; CLSCT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
KW Lectin; Hydroxylation; Glycoprotein; Mannose-binding; Membrane;
KW Collagen; Repeat; Calcium; Signal.
FT SIGNAL 1 20
FT CHAIN 21 371 CONGLUTININ.
FT DOMAIN 46 216 COLLAGEN-LIKE.
FT DOMAIN 273 371 C-TYPE LECTIN (SHORT FORM).
FT MOD_RES 63 63 HYDROXYLATION.
FT MOD_RES 87 87 HYDROXYLATION.
FT MOD_RES 99 99 HYDROXYLATION.
FT MOD_RES 135 135 HYDROXYLATION.
FT MOD_RES 141 141 HYDROXYLATION.
FT MOD_RES 159 159 HYDROXYLATION.
FT MOD_RES 162 162 HYDROXYLATION.
FT MOD_RES 198 198 HYDROXYLATION.
FT SITE 201 203 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 275 369 BY SIMILARITY.
FT DISULFID 347 361 BY SIMILARITY.
FT CARBOHYD 337 337 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 173 173 R -> H (IN REF. 2 AND 3).
FT CONFLICT 210 210 K -> S (IN REF. 6).
FT CONFLICT 218 218 V -> A (IN REF. 2).
FT CONFLICT 272 272 E -> V (IN REF. 2).
SQ SEQUENCE 371 AA; 37994 MW; 867BB41992544B1P CRC64;

Alignment Scores:
Pred. No.: 4,41e-29 Length: 371
Score: 441.50 Matches: 91
Percent Similarity: 56.77% Conservative: 39
Best Local Similarity: 39.74% Mismatches: 74
Query Match: 32.34% Indels: 25
DB: 1 Gaps: 6

US-10-054-536-2 (1-747) x CONG_BOVIN (1-371)
Qy 118 TCTCCAGGCATCAACGGCTTCCAGGCAAGATGGCGTGATGCACCAAGGGAGAAAG 177
Db 146 AlaProGlyTleGlnGlyPheProGlyPro-----SerGlyLeuLysGlyGluLys 162
Qy 178 GGGGAACACAGGCCAA-----GGGCTCAGAGCTTACAGGCCCCCTGGGAAG 225
Db 163 GlyAlaProGlyGluThrGlyAlaProGlyArgAlaGlyValThrGlyProSerGlyAla 182
Qy 226 TTGGGGCTCCAGGAATCCAGGCTTCTGGGTACACAGGACCAAGGGCAAAAGGA 285
Db 183 IleGlyProGlnGlyProSerGlyAlaArgGlyProGlyLeuLysGlyAspArgGly 202
Qy 286 GACCTGGAAAAGT---CCGGATGGTATAGTAGCTGGTGGC----- 327
Db 203 AspProGlyGluThrGlyAlaLysGlyGluSerGlyLeuAlaGluValAsnAlaLeuLys 222
Qy 328 -----TCAGAAAGAAAGAACTCTGCAAAACAGAAAATGGCAGCATC 366
Db 223 GlnArgValThrIleLeuAspGlyHisLeuArgArgPheGlnAsnAlaPheSerGlnTyr 242
Qy 367 AAAAAGTGGCTGACCTTCTCTGGGCAAAACAAGTTGGGAACAAAGTTCTCTGACCAAT 426
Db 243 LysLysAlaValLeuPheProAspGlyGlnAlaValGlyGluLysIlePheLysThrAla 262
Qy 427 GGTGAATAATGACCTTTGAAAAAGTCAAGGCTTGTGTGTCAGAGTTCCAGGCTCTGTG 486
Db 263 GlyAlaValLysSerTyrSerAspAlaGluGlnLeuCysArgGlnAlaLysGlyClnLeu 282


```

Db      164 LysGlyGluArgGlyAlaProGlyGluGlnGlyAlaProGlyAsnAlaGlyAlaGly 183
Qy      309 -----
Db      184 ProAlaGlyProAlaGlyProGlnGlyAlaProGlySerArgGlyProProGlyLeuIlys 203
Qy      310 -----CATAGT 315
Db      204 GlyAspArgGlyAlaProGlyAspArgGlyIleIysGlyGluSerGlyLeuProAspSer 223
Qy      316 AGCTCGTGCTCCTCAGAAAGAAAGCTCTGCAAAACAGAAATCGCAGCTATC----- 366
Db      224 AlaAlaLeuArgGlnGlnMetGluAlaLeuAsnGlyLysLeuGlnArgLeuGluAla 243
Qy      367 -----AAAAAGTGCTGACCTCTCTCTCTGCGCAACCAAGTTGGGAACAAGTTC 414
Db      244 PheSerArgTyrLysAlaAlaLeuPheProAspGlyGlnSerValGlyAspIysIle 263
Qy      415 TTCCTGACCAATGTGAAATAATGACCTTTCGAAAAGTGAAGCGCTTGTCGTCAGTTC 474
Db      264 PheArgAlaAlaAsnSerGluGluProPheGluAspAlaLysGluMetCysArgGlnAla 283
Qy      475 CAGGCTCTGTGGCCACCCCGAGGAATCTGCAGAGAATGGAGCCATTCAGAATCTCATC 534
Db      284 GlyGlyGlnLeuAlaSerProArgSerAlaThrGluAsnAlaAlaValGlnGlnLeuVal 303
Qy      535 -----AGGAGGAGCCTCTCTGGGCATCACTGATGAGACAGACAGAGGCGAGTTT 585
Db      304 ThrAlaHisSerLysAlaAlaPheLeuSerMetThrAspValGlyThrGluGlyIysPhe 323
Qy      586 GTGATCTGCAGAGAATAGACTGACCTACCACTGCAACAACTGGAAGGGTGAACCCCAACAT 645
Db      324 ThrTyrProThrGlyGluAlaLeuValTyrSerAsnTyrAlaProGlyGluProAsnAsn 343
Qy      646 GCTGTTCTGTAGAGATTGTGTATTGCTATGAAATGCGCAGTGAATGAGTCCCTCCC 705
Db      344 AsnGlyGlyAlaGluAsnCysValGluIlePheThrAsnGlyGlnTyrAsnAspIysAla 363
Qy      706 TGCTCCACCTCCCTCTGGCGTGTGAGTTC 738
Db      364 CysGlyGluGlnArgLeuValIleCysGluPhe 374

RESULT 11
PSPD_MOUSE
ID_PSPD_MOUSE STANDARD; PRT; 374 AA.
AC PS0404;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
GN SFTPD OR SFTP4
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=96094460; PubMed=7499852;
RA Motwani M., White R.A., Guo N., Dowler L.L., Tauber A.I., Sastry K.N.;
RT "Mouse surfactant protein-D. cDNA cloning, characterization, and gene
RL localization to chromosome 14."
RL J. Immunol. 155:5671-5677 (1995).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=99244602; PubMed=10226065;
RA Lawson P.R., Perkins V.C., Holmskov U., Reid K.B.;
RT "Genomic organization of the mouse gene for lung surfactant protein
RL D."
RL Am. J. Respir. Cell Mol. Biol. 20:953-963 (1999).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RA Fisher J.H., Sheftelyevich V.V.;
RT "Surfactant protein-D regulates surfactant phospholipid homeostasis in
RL vivo."
RN Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Mariani K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson K.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RP human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Contributes to the lung's defense against inhaled
CC microorganisms. Binds strongly maltose residues and to a lesser
CC extent other alpha-glucosyl moieties. It could participate in the
CC extracellular reorganization or turnover of pulmonary surfactant.
CC -!- SUBUNIT: Oligomeric complex of 4 set of homotrimers.
CC -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC protein. There are 4 surfactant-associated protein: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC hydrophobic proteins (SP-B and SP-C).
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L40156; AAA92021.1; --
CC EMBL; AF047742; AAD31380.1; --
CC EMBL; AF047741; AAD31380.1; JOINED.
CC EMBL; AF192134; AAF15277.1; --
CC EMBL; BC003705; AAH03705.1; --
CC HSSP; P35247; 1B08.
CC MGD; MGI:109515; Sftpd.
CC InterPro; IPR008160; Collagen.
CC InterPro; IPR001304; Lectin_C.
CC Pfam; PF01391; Collagen; 3.
CC Pfam; PF00059; Lectin_C; 1.
CC SMART; SM00034; CLECT; 1.
CC PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
CC PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
KW Glycoprotein; Calcium; Surface Film; Gaseous exchange; Hydroxylation;
KW Signal; Lectin; Collagen; Repeat; Coiled coil.
FT SIGNAL 1 19
FT CHAIN 20 374
FT DOMAIN 45 221
FT DOMAIN 222 253
FT DOMAIN 278 374
FT DISULFID 280 372
FT DISULFID 350 364

```



```
FT DISULFID 275 367 BY SIMILARITY.
FT CARBOHYD 345 359 BY SIMILARITY.
FT MOD_RES 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MOD_RES 78 87 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 87 87 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 96 96 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 99 99 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 165 165 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 171 171 HYDROXYLATION (BY SIMILARITY).
SQ SEQUENCE 369 AA; 37361 MW; 07D88B24E0ABE2E3 CRC64;

Alignment Scores:
Pred. No.: 6,47e-27 Length: 369
Score: 415.50 Matches: 87
Percent Similarity: 52.23% Conservatives: 30
Best Local Similarity: 38.84% Mismatches: 90
Query Match: 30.44% Indels: 17
DB: 1 Gaps: 4

US-10-054-536-2 (1-747) x PSPD_BOVIN (1-369)
QY 118 TCTCAGCATCAAGGCTTCCAGGCAAGATGGCGTGATGGCACCACGAGGAAAG 177
Dy 146 AlAProGlyIleGlnGlySerProGlyProAlaGlyLeuLysGlyAlaPro 165
QY 178 GGGGAACCAAGCC---CAAGGGCTCAGAGGCTTACAGGGCCCCCTGGAACTTGGGGCT 234
Dy 166 GlyGluProGlyAlaProGlyArgAlaGlyAlaProGlyProAlaGlyAlaIleGlyPro 185
QY 235 CCAGGAATCCAGGCGCTTCTGGGTACCCAGGACCAAGAGGCCCAAGAGAGAGAGAGAG 294
Dy 186 GlnGlyProSerGlyAlaArgGlyProGlyLeuLysGlyAspArgGlyThrProGly 205
QY 295 ---AAAGTCGGATGTGTAGTAGCTGCT----- 324
Dy 206 GluArgGlyAlaLysGlyGluSerGlyLeuAlaGluValAsnAlaLeuArgGlnArgVal 225
QY 325 -----GCCTCAGAAAGAAAGCTCTGCAACAGAGAAATGGCACGTATCAAAAGTGG 375
Dy 226 GlyIleLeuGluGlyGlnLeuGlnArgLeuGlnAsnAlaPheSerGlnTyrLysLysAla 245
QY 376 CTGACCTTCTCTGGGCAACAAAGTTGGGAACAAAGTTCTTCTGACCAATGGTGAATA 435
Dy 246 MetLeuPheProAsnGlyArgSerValGlyGluLysIlePheLysThrValGlySerGlu 265
QY 436 ATGACCTTGAAGAAAGTGAAGCCCTTGTGTCTCAAGTTCAGGCTCTGTGGCCACCC 495
Dy 266 LysThrPheGlnAspAlaGlnGlnIleCysThrGlnAlaGlyGlnLeuProSerPro 285
QY 496 AGGAATGCTCAGAGAATGGAGCCATTCAAGATCTCATC-----AAGGAGGAGAGCC 546
Dy 286 ArgSerGlyAlaGluAsnGluAlaLeuThrGlnLeuAlaThrAlaGlnAsnLysAlaAla 305
QY 547 TTCCTGGGCATCTAGATGAGAAGACAGAGGCGAGTTTGTGATCTGACAGAAATAGA 606
Dy 306 PheLeuSerMetSerAspThrArgLysGluGlyThrPheIleTyrProThrGlyGluPro 325
QY 607 CTGACCTACACAACTGAACGAGGCTGACCCCAACATGCTGCTGTGATGAGATGT 666
Dy 326 LeuValTyrSerAsnTyrPalaProGlnGluProAsnAsnAspGlyGlySerGluAsnCys 345
QY 667 GTATTGCTACTGAAATGGCCAGTGAATGACCTGCCCTGCTCCACCTCCCATCTGCCC 726
Dy 346 ValGluIlePheProAsnGlyLysTyrAsnAspLysValCysGlyGluGlnArgLeuVal 365
QY 727 GTCGTGAGTTC 738
Dy 366 IleCysGluPhe 369

RESULT 15
PSPA_CAVPO
ID -PSPA_CAVPO STANDARD; PRT; 247 AA.
AC P50403;
```

```
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)
DE (PSPAP).
DE SFTPA1 OR SFTPA OR SFTPA1.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley; TISSUE=Lung;
EX MEDLINE=98018900; PubMed=9357868;
RA Yuan H.T., Gowan S., Kelly F.O., Bingle C.D.;
RT "Cloning of guinea pig surfactant protein A defines a distinct
RL cellular distribution pattern within the lung.";
RL Am. J. Physiol. 273:L900-L906(1997).
CC -!- FUNCTION: In presence of calcium ions, PSAP binds to surfactant
CC phospholipids and contributes to lower the surface tension at the
CC air-liquid interface in the alveoli of the mammalian lung and is
CC essential for normal respiration.
CC -!- SUBUNIT: Oligomeric complex of 6 set of homotrimers.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC protein. There are 4 surfactant-associated protein: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-B and SP-C).
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40869; AAB82952.1; -.
DR HSPSP; P22897; LEGG.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01391; Collagen; 1.
DR Pfam; PF00059; lectin_C; 1.
DR ProDom; PD000007; Clg_helix; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C TYPE LECTIN 1; 1.
DR PROSITE; PS00041; C TYPE LECTIN 2; 1.
DR GlycoProtein; Calcium; Surface Film; Gaseous exchange; Hydroxylation;
DR Signal; Lectin; Collagen; Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 247 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT A.
FT DOMAIN 27 99 COLLAGEN-LIKE.
FT DOMAIN 152 245 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 154 245 BY SIMILARITY.
FT DISULFID 223 237 BY SIMILARITY.
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 247 AA; 26104 MW; DIBC6270EEFC932 CRC64;

Alignment Scores:
Pred. No.: 4,04e-26 Length: 247
Score: 405.50 Matches: 98
Percent Similarity: 52.71% Conservatives: 38
Best Local Similarity: 37.98% Mismatches: 89
Query Match: 29.71% Indels: 33
DB: 1 Gaps: 10

US-10-054-536-2 (1-747) x PSPA_CAVPO (1-247)
```


B/gnk

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 21, 2004, 16:39:39 ; Search time 3234 Seconds
(without alignments)
10011.517 Million cell updates/sec
Title: US-10-054-536-2
Perfect score: 747
Sequence: 1 atgtccctgtttccatcact.....tctgtgagttccctatctga 747

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pin.*
- 35: em.htg_rod.*
- 36: em.htg_man.*
- 37: em.htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	747	100.0	747	6	E27637	E27637 Recombinant
2	747	100.0	900	6	AR182149	AR182149 Sequence
3	747	100.0	1632	9	HOSA16576	Y16576 Homo sapien
4	747	100.0	1638	9	HOSA16577	Y16577 Homo sapien
5	747	100.0	1638	9	HOSA16580	Y16580 Homo sapien
6	747	100.0	1638	9	HOSA16581	Y16581 Homo sapien
7	747	100.0	3605	6	E27636	E27636 Recombinant
8	747	100.0	3605	6	AX411061	AX411061 Sequence
9	747	100.0	3605	9	HSMBPC	X15422 Human mRNA
10	745.4	99.8	1632	9	HOSA16578	Y16578 Homo sapien
11	745.4	99.8	1638	9	HOSA16579	Y16579 Homo sapien
12	745.4	99.8	1638	9	HOSA16582	Y16582 Homo sapien
13	743.8	99.6	776	9	AF360391	AF360391 Homo sapi
14	635.8	85.1	1188	6	I09220	I09220 Sequence 1
15	587.8	78.7	805	9	MACMBPC	L43911 Macaca mula
16	444.6	59.5	1409	4	D73408	D73408 Bos taurus
17	406.6	54.4	723	4	AF164576	AF164576 Sus scrof
18	395	52.9	1010	6	E37364	E37364 Swine serum
19	381.4	51.1	799	9	MACMBPA	L43912 Macaca mula
20	374.8	50.2	1116	10	S42294	S42294 mannose-bin
21	374	50.1	1254	9	HSMBPID	X15957 H.sapiens M
22	374	50.1	3336	9	HSMBPCA3	AF080510 Homo sapi
23	373.2	50.0	1069	10	MUSRRFA	D11440 Mus musculu
24	373.2	50.0	1098	10	BC010760	BC010760 Mus muscu
25	372.4	49.9	102532	9	AL731550	AL731550 Human DNA
26	370.8	49.6	302859	2	AC044785	AC044785 Homo sapi
27	366.2	49.0	874	10	RNMBER	X05023 Rat mRNA fo
28	364.6	48.8	1037	6	AX402034	AX402034 Sequence
29	364.6	48.8	1037	10	RATMABPC	M14103 Rat mannose
30	287.6	38.5	1068	6	AX827373	AX827373 Sequence
31	287.6	38.5	1068	10	AF080507	AF080507 Rattus sp
32	286	38.3	866	10	MUSRRFB	D11441 Mus musculu
33	286	38.3	894	10	BC021762	BC021762 Mus muscu
34	284.4	38.1	943	6	E63781	E63781 Serum manna
35	284.4	38.1	943	10	S42292	S42292 mannose-bin
36	280.6	37.6	1423	9	AF019382	AF019382 Homo sapi
37	206	27.6	141678	10	AC102722	AC102722 Mus muscu
38	204.4	27.4	1211	6	AR182169	AR182169 Sequence
39	204.4	27.4	1211	10	MMMBI16	U09016 Mus musculu
40	204.4	27.4	223274	2	AC112443	AC112443 Rattus no
41	204.4	27.4	270783	2	AC105489	AC105489 Rattus no
42	193.8	25.9	190092	2	AC120502	AC120502 Lemur cat
43	192	25.7	134411	9	AL512662	AL512662 Human DNA
44	188	25.2	1110	9	HSMBPIA	X15954 H.sapiens M
45	188	25.2	1802	9	HSMBPCAL	AF080508 Homo sapi

ALIGNMENTS

RESULT 1
E27637
LOCUS E27637
DEFINITION Recombinant human mannan binding protein and process for producing the same.
ACCESSION E27637
VERSION E27637.1 GI:13018239
KEYWORDS JP 1999206378-A/2.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 747)
AUTHORS Nobutaka,W.
TITLE Recombinant human mannan binding protein and process for producing the same

747 bp DNA linear PAT 18-JUN-2001

RESULT 10	HOSAL16578	Homo sapiens gene encoding mannan/mannose-binding protein, variant LYQC.	1632 bp	DNA	linear	PRI 17-SEP-1999
LOCUS	Y16578	Y16578.1	GI:59111793			
DEFINITION	manose-binding lectin; mbl gene.					
ACCESSION	Y16578					
VERSION	Y16578.1					
KEYWORDS	manose-binding lectin; mbl gene.					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1	Madsen, H.O., Satz, M.L., Høgh, B., Svejgaard, A. and Garred, P. Different molecular events result in low protein levels of mannan-binding lectin in populations from southeast Africa and South America				
AUTHORS	Madsen, H.O.					
TITLE	Direct Submission					
JOURNAL	Submitted (16-FEB-1998) H.O. Madsen, Department of Clinical Immunology, Section 7631, National University Hospital, Tagensvej 20, DK-2200 Copenhagen, DENMARK					
COMMENT	Related sequences X15954, X15955, X15956, X15422.					
FEATURES	Location/Qualifiers					
source	1..1632					
	/organism="Homo sapiens"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:9606"					
	/chromosome="10"					
	/map="10q11.2-q21"					
	/note="MBL haplotype LYQC"					
variation	273					
	/note="polymorphism in different MBL haplotypes"					
variation	396					
	/note="polymorphism in different MBL haplotypes"					
variation	474					
	/note="polymorphism in different MBL haplotypes"					
variation	487					
	/note="polymorphism in different MBL haplotypes"					
variation	596					
	/note="polymorphism in different MBL haplotypes"					
gene	739..1632					
	/gene="MBL"					
CAAT_signal	739..744					
	/gene="MBL"					
variation	747					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
TATA_signal	780..785					
mRNA	817..>1632					
	/gene="MBL"					
variation	820					
	/product="mannose-binding lectin"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
CDS	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					
	/gene="MBL"					
	/codon_start=1					
	/product="mannose-binding lectin"					
	/protein_id="CAB56045.1"					
	/db_xref="GI:59111794"					
	/db_xref="GOA:P11226"					
	/db_xref="SWISS-PROT:P11226"					
	/translation="MSLFFSLPLLLLSMVAASYSERVTCEDAQKTCPAVIACSSPGIN GPGKDRGDTKEEGEPGLGQPGKLGPPGNPGSPGPGKQKPGKSPD GDSLSAASRKALQETEMARIKKWLTFSLGKQVGNKFFLTNGEIMTFEKVKALCVKFOA SVATPRNAEANGAIQNLKEEAFGLITDEKTEQGFVDLTGNRLTYTNWNEGEPPNAGS					
	/gene="MBL"					
	/note="polymorphism in different MBL haplotypes"					
	886..1632					

AF360991.1		GI:14030459
VERSION	Homo sapiens (human)	
KEYWORDS	Homo sapiens	
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	Chen,Z., Zhu,X. and Xie,P. Cloning and sequencing of mannan-binding lectin cDNA of Chinese Milanixue Zazhi 15, 83-86 (1999) 2 (bases 1 to 776)	
AUTHORS	Chen,Z., Zhu,X. and Xie,P.	
JOURNAL	Direct Submission	
TITLE	Submitted (13-MAR-2001) Immunology, First Military Medical University, Tonghe, Guangzhou, Guangdong 510515, P.R. of China	
JOURNAL	Location/Qualifiers	
FEATURES	1..776 /source	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/chromosome="10"	
	/map="10q11.2-q21"	
	/cell_type="hepatocyte"	
	/tissue_type="liver"	
gene	1..776_MBL	
CDS	17..763	
	/gene="MBL"	
	/notes="activates complement system; serum protein"	
	/codon_start=1	
	/product="mannan-binding lectin MBL precursor"	
	/protein_id="AAK52907.1"	
	/gb_xref="GI:14030460"	
	/translation="MSLPFLPLLLSMVAASYSETVACEDAQKTCPAVIACSPPGIN GFGPKDGRGTGEKGEPQGLQGPGKPNGPNPSPGSPPKGGDKDPKPSFD GPSLAASERKALOTEMARIKKWLTFSLGVKNKFELTNGEIMTEFKVKALCVKFOA SVATPRNAENGAIQNLIKKEEAFLGITBEKTEGGQVDLTGNRLFYTNWEGEPNAGS BDCVLILKNGWNVPCTSHLVACEFPI"	
sig_peptide	17..76	
mat_peptide	77..760	
	/gene="MBL"	
	/product="mannan-binding lectin MBL"	
variation	86	
	/gene="MBL"	
	/note="compared to Homo sapiens mannan-binding lectin encoded by GenBank Accession Number XI5422; Ala to Thr substitution"	
	/replace="A"	
variation	571	
	/gene="MBL"	
	/note="compared to Homo sapiens mannan-binding lectin encoded by GenBank Accession Number XI5422; does not alter sequence of predicted protein"	
	/replace="C"	
ORIGIN		
Query Match	99.6%; Score 743.8; DB 9; Length 776;	
Best Local Similarity	99.7%; Pred.No. 2e-208;	
Matches	745; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Qy	1 ATGTCCCTTTTCACACTCCCTCTCCTTCTCTCGATGATGGTGCGAGCGCTTACTCA 60	
Db		
	17 ATGTCCCTGTTTCCCATCACTCCCTCTCCTTCTCTCGATGATGGTGCGAGCGCTTACTCA 76	
Qy	61 GAAACTGTGACCTGTGAGGATGCCAAAAGACCTGCCTGCAGTGATTGCGCTTAGCTCT 120	
Db		
	77 GAAACTGTGCGCTGTGAGGATGCCCAAAGACCTGCCTGCAGTGATTGCGCTTAGCTCT 136	
Qy	121 CAGGATCAACGGCTTCCAGCAAGAATGGCGTGTATGGCACCAAGGGAAGAAAGGG 180	
Db		
	137 CCAGGATCAACGGCTTCCAGCAAGAATGGCGTGTATGGCACCAAGGGAAGAAAGGG 196	

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 18, 2004, 20:52:46 ; Search time 56.5 Seconds
(without alignments)
7471.256 Million cell updates/sec

Title: US-10-054-536-2

Perfect score: 1365

Sequence: 1 atgtccctgtttccatcact.....tctgtgagttccctatctga 747

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xl
-Q=/cgn2.1/USPTO.spool p/US10054536/runat_18062004_173454_13562/app_query.fasta_1.903
-DB=A Geneseq 29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10054536 @CGN 1.1.81/runat_18062004_173454_13562 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq 29Jan04.*

1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1321	96.8	248	2 AAY29485	Human man
2	1321	96.8	248	4 AAB36578	Human man
3	1321	96.8	248	5 AAU75574	Human man
4	1321	96.8	248	7 ADE87458	248 amino
5	1316	96.4	248	2 AAR45005	Human Man
6	1309	95.9	248	3 AAY77988	Human MBP
7	1228	90.0	228	2 AAY03769	Mannan-bi
8	1137	83.3	247	1 AAP94680	Amino aci
9	806.5	59.1	249	3 AAB15401	Pig serum
10	727	53.3	409	4 AAM79745	Human pro

11	474	34.7	282	6 ABR42241	Human sur
12	454	33.3	375	3 AAY77990	Human SP-
13	454	33.3	375	5 AAU76473	Human lun
14	454	33.3	375	6 ABR44021	Human sur
15	442.5	32.4	351	2 AAM18780	Bovine co
16	442.5	32.4	351	2 AAM13672	Bovine co
17	441	32.3	374	5 AAU76471	Rat lung
18	440.5	32.3	371	2 AAR75642	Bovine co
19	408	29.9	126	5 AEG80907	Rat MBP p
20	407	29.8	131	5 AEG80919	Rat PrMBP
21	404	29.6	248	5 AAU76476	Antioxi
22	402	29.5	249	7 ADD48682	Rat Prote
23	401	29.4	335	2 AAR12222	Conglutin
24	398.5	29.2	248	3 AAY77989	Human SP-
25	398.5	29.2	248	5 AAU76468	Human lun
26	398.5	29.2	249	7 ADD48683	Human Pro
27	397	29.1	116	4 AAM78761	Human pro
28	396.5	29.0	243	7 ADD49123	Human NOV
29	394.5	28.9	248	1 AAP60441	Plasmid p
30	394.5	28.9	248	1 AAP60665	Sequence
31	394.5	28.9	248	1 AAP80694	Sequence
32	394.5	28.9	248	2 AAR04215	Human 32K
33	394.5	28.9	248	2 AAR06331	Human alv
34	394.5	28.9	248	7 ADD49127	Human NOV
35	393.5	28.8	248	2 AAR04216	Human 32K
36	393.5	28.8	271	2 AAR04212	Human 32K
37	393.5	28.8	271	2 AAR04217	Human 32K
38	390.5	28.6	248	1 AAP70662	35kd pulm
39	389.5	28.5	259	3 AAB58135	Lung canc
40	387.5	28.4	248	2 AAR05091	Vector PS
41	386.5	28.3	248	2 AAR05092	Product o
42	385.5	28.2	248	1 AAP60442	Plasmid p
43	385.5	28.2	248	1 AAP60666	Genomic s
44	385.5	28.2	248	1 AAP70663	35kd pulm
45	385.5	28.2	271	1 AAP60661	Genomic s

ALIGNMENTS

RESULT 1

AAY29485

ID AAY29485 standard; protein; 248 AA.

AC AAY29485;

DT 11-OCT-1999 (first entry)

DE

Human mannan-binding protein.

KW

Human; mannan-binding protein; hMBP; recombinant; inhibition; infection;

rhMBP; haemagglutination; influenza; HIV.

XX

Homo sapiens.

XX

Key Location/Qualifiers

FT Peptide 1..20

FT Protein 21..248

FT /label= mannan-binding_protein

XX WO9937676-A1.

XX 29-JUL-1999.

XX 23-JUL-1998; 98WO-JP003311.

XX 23-JAN-1998; 98JP-00011864.

XX (FUSO) FUSO PHARM IND LTD.

XX Wakamiya N;

XX WPI; 1999-469114/39.

CC particular infections which are seen in connection with implantation
 CC and/or transplantation of organs and diseases such as chronic
 CC inflammatory demyelinating polyneuropathy, multiple sclerosis, epilepsy,
 CC myasthenia gravis, Eaton-Lambert's syndrome, rheumatoid arthritis,
 CC systemic lupus erythematosus, vasculitis, Wegner's granulomatosis,
 CC Sjogren's syndrome, autoimmune neutropenia, Crohn's disease, colitis
 CC ulcerous, asthma, septic shock syndrome, psoriasis, toxic shock syndrome,
 CC diabetes, sinusitis endocarditis, atherosclerosis, primary
 CC hypo/agammaglobulinemia including common variable immunodeficiency,
 CC Wiskott-Aldrich syndrome and severe combined immunodeficiency (SCID),
 CC Kawasaki's disease, and Guillain-Barre's syndrome. The present sequence
 CC represents human MBL, which is given in the exemplification of the
 CC present invention
 XX
 SQ Sequence 248 AA;

Alignment Scores:
 Pred. No.: 5,48e-115 Length: 248
 Score: 1321.00 Matches: 248
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.78% Indels: 0
 DB: 4 Gaps: 0

US-10-054-536-2 (1-747) x AAB36578 (1-248)

QY 1 ATGTCCTGTTTCCATCACTCCCTCTCCCTTCCTGAGTATGCTGCGAGCGTCTTACTCA 60
 DB 1 MetSerLeuPheProSerLeuProLeuLeuLeuLeuSerMetValAlaAlaSerTyrSer 20
 QY 61 GAAACGTGACCTGAGGATGCCCAAAAGACCTGCCCTGCAGTGTGCTGTTAGCTCT 120
 DB 21 GluThrValThrCysGluAspAlaGlnLysThrCysProAlaValLeuAlaCysSerSer 40
 QY 121 CAAGCATCAACGGCTTCCAGGCAAGATGGCGCTGTATGGCACCAAGGAGCAAAAGGGG 180
 DB 41 ProGlyIleAsnGlyPheProGlyLysAspGlyArgAspGlyThrLysGlyGluLysGly 60
 QY 181 GAACGAGCAAGGGCTCAGAGCTTACAGGCGCCCTGGAAGTTGGGGCTCCAGGA 240
 DB 61 GluProGlyGlnGlyLeuArgGlyLeuGlnGlyProProGlyLysLeuGlyProGly 80
 QY 241 AATCCAGGCGCTTCTGGGTCCACGAGCAACAAAGGCGCAAAAGGACCTCGAAAAAGCT 300
 DB 81 AsnProGlyProSerGlySerProGlyProLysGlyGlnLysGlyAspProGlyLysSer 100
 QY 301 CCGATGTGTATAGTACCTGCTGCTCAGAAAGAAAGCTCTGCAACAGAAATGGCA 360
 DB 101 ProAspGlyAspSerSerLeuAlaAlaSerGluArgLysAlaLeuGlnThrGluMetAla 120
 QY 361 CGTATCAAAAGTGGTGCCTTCTCTGGGCAACCAAGTTGGGAACAGTCTTCTCTG 420
 DB 121 ArgIleLysLysTrpLeuThrPheSerLeuGlyLysGlnValGlyAsnLysPheLeu 140
 QY 421 ACCAATGTGTAATATGACCTTGAAGAAGTGAAGCGCTTGTCTCAAGTTCACAGGC 480
 DB 141 ThrAsnGlyGluIleMetThrPheGluLysValLysAlaLeuCysValLysPheGlnAla 160
 QY 481 TCTGTGGCCACCCCGAGGAATGCTCAGAGATGGAGCCATTGAGAAATCTCATCAGGAG 540
 DB 161 SerValAlaThrProArgAsnAlaAlaGluAsnGlyAlaIleGlnAsnLeuLysGlu 180
 QY 541 GAAGCCTTCTGGGCATCATCTGATGAGAGACAGAAAGGGGAGTTGTGATGATCAGAGA 600
 DB 181 GluAlaPheLeuGlyIleThrAspGluLysThrGluGlyGlnPheValAspLeuThrGly 200
 QY 601 AATAGACTCACCTACCAACTGGACAGGGGTGAACCCCAACAACTCTGTTCTGTATGAA 660
 DB 201 AsnArgLeuThrTyrThrAsnTrpAsnGlnGlyGluProAsnAsnAlaGlySerAspGlu 220
 QY 661 GATTGTGATTGTACTGAAAAATGCCAGTGAATGAGCTGCCCTGTCTCCACCTCCCAT 720
 DB 221 AspCysValLeuLeuLysAsnGlyGlnTrpAsnAspValProCysSerThrSerHis 240

QY 721 CTGGCGGTCTGTGAGTTCCTTATC 744
 DB 241 LeuAlaValCysGluPheProIle 248
 RESULT 3
 AAU75574
 ID AAU75574 standard; protein; 248 AA.
 XX
 AC AAU75574;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human mannose-binding protein (MBP).
 XX
 KW Human; mannose-binding protein; MBP; methylotrophic yeast strain;
 KW protein disulphide isomerase; PDI; heat shock protein 47; hsp47;
 KW propyl-4-hydroxylase; P4H; pathogenic organism disposal; biocide.
 OS Homo sapiens.
 XX
 PN US6337193-B1.
 XX
 PD 08-JAN-2002.
 XX
 PF 24-NOV-1998; 98US-00198603.
 XX
 PR 24-NOV-1998; 98US-00198603.
 XX
 PA (APTA-) APTAGEN INC.
 XX
 PI Tully RE, Caltagirone GT, Moyer SS, Ronning MT;
 XX
 DR WPI; 2002-163238/21.
 DR N-PSDB; ABK14771.

Methylotrophic yeast strain, useful for producing mannose-binding protein, comprises DNA molecules encoding the protein, protein disulphide isomerase, heat shock protein 47 and propyl-4-hydroxylase.
 Example 1; Fig 1; 30pp; English.
 The invention relates to a methylotrophic yeast strain (I) comprising a DNA molecule encoding mannose-binding protein (MBP), protein disulphide isomerase (PDI), heat shock protein 47 (hsp47), and propyl-4-hydroxylase (P4H), where upon culturing (I) produces the MBP, PDI, hsp47 and P4H. (I) is useful for producing MBP comprising culturing (I) under conditions suitable for the secretion of MBP by the yeast where MBP is utilised in disposal of pathogenic organisms by opsonising pathogen or activating complement cascade. The methods utilising (I) result in high yields of MBP without the use of foetal calf serum, in a cost-effective manner. The present sequence relates to the amino acid sequence of human mannose-binding protein

Alignment Scores:
 Pred. No.: 5,48e-115 Length: 248
 Score: 1321.00 Matches: 248
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.78% Indels: 0
 DB: 5 Gaps: 0

SQ Sequence 248 AA;

US-10-054-536-2 (1-747) x AAU75574 (1-248)

QY 1 ATGTCCTGTTTCCATCACTCCCTCTCCCTTCCTGAGTATGCTGCGAGCGTCTTACTCA 60
 DB 1 MetSerLeuPheProSerLeuProLeuLeuLeuSerMetValAlaAlaSerTyrSer 20
 QY 61 GAAACGTGACCTGTGAGGATGCCCAAAAGACCTGCCCTGCAGTGTGCTGTTAGCTCT 120
 DB 21 GluThrValThrCysGluAspAlaGlnLysThrCysProAlaValLeuAlaCysSerSer 40

CC plasmid or viral vector (especially in a vaccinia or adenovirus vector).
CC The MBP may be human MBP, or derived from it by addition, deletion and/or
CC substitution of one or more amino acid residues. The agents containing
CC the MBP can be used for effective treatment of malignant tumours
CC (including solid tumours) by direct application to the tumour. The
CC present sequence represents a MBP sequence
XX
SQ Sequence 228 AA;

Alignment Scores:
Pred. No.: 2,79e-106 Length: 228
Score: 1228.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.96% Indels: 0
DB: 2 Gaps: 0

US-10-054-536-2 (1-747) x AAY03769 (1-228)

QY 61 GAACTGTGACCTGTGAGGATCCCAAGACCTGCCCTGCAGTGATTGCTGTAGCTCT 120
DB 1 GluThrValThrCysGluAspAlaGlnLysThrCysProAlaValIleAlaCysSer 20
QY 121 CCAGGCATCAACGGCTTCCACGCCAAGATGGCGCTGATGGCACCAAGGGAGAAAGGG 180
DB 21 ProGlyIleAsnGlyPheProGlyLysAspGlyArgAspGlyThrLysGlyGluLysGly 40
QY 181 GAACGAGCCAAAGGCTCAGAGCTTACAGGCCGCCCTGGAAAGTTGGGGCTCCAGGA 240
DB 41 GluProGlyGlnGlyLeuArgGlyLeuGlnGlyProGlyLysLeuGlyProGly 60
QY 241 AATCAGGCTTCTGGGTACAGACCAAGGCCAAGGCCAAGGCCAAGGCCAAGGCCAAG 300
DB 61 AsnProGlyProSerGlySerProGlyProLysGlyGlnLysGlyAspProGlyLysSer 80
QY 301 CCGGATGGTGATAGTACCTGCTGCTCAGAAAGAAAGCTCTGCAACACAAATGGCA 360
DB 81 ProAspGlyAspSerSerLeuAlaAlaSerGluArgLysAlaLeuGlnThrGluMetAla 100
QY 361 CGTATCAAAAAGTGGCTGACCTTCTCTCGGCCAACCAAGTTGGGAACAAGTTCTCTCG 420
DB 101 ArgIleLysLysTriLeuThrPheSerLeuGlyLysGlnValGlyAsnLysPheLeu 120
QY 421 ACCAATGGTGAATATATGACCTTTGAAAGTGAAGGCCCTTGTGTGTCAGTTCAGGCC 480
DB 121 ThrAsnGlyGluIleMetThrPheGluLysValLysAlaLeuCysValLysPheGlnAla 140
QY 481 TCTGTGGCCACCCAGGAATGCTGAGAGATGGAGCCATTGAGAAATCTCATCAAGGAG 540
DB 141 SerValAlaThrProArgAsnAlaAlaGluAsnGlyAlaIleGlnAsnLeuLysGlu 160
QY 541 GAAGCCTTCTGGGCATCACTGATGAGACAGAGAGGCCAGTTTGTGATCTGACAGA 600
DB 161 GluAlaPheLeuGlyIleThrAspGluLysThrGluGlyGlnPheValAspLeuThrGly 180
QY 601 AATAGACTGACCTACACAACTGGAACGAGGTGAACCCCAACAAATGCTGTTCTGATGA 660
DB 181 AsnArgLeuThrTyThrAsnTrpAsnGluGlyGluProAsnAsnAlaGlySerAspGlu 200
QY 661 GATTGTGATTCTACTGAAAATGCCCAGTGAATGAGCTGCCCTGCTCCACCTCCCAT 720
DB 201 AspCysValLeuLeuLeuLysAsnGlyGlnTrpAsnAspValProCysSerThrSerHis 220
QY 721 CTGGCCGCTGTGAGTTCCTATC 744
DB 221 LeuAlaValCysGluPheProIle 228

RESULT 8
AAP94680
ID AAP94680 standard; protein; 247 AA.
XX
AC AAP94680;
XX

DT 25-MAR-2003 (revised)
DT 29-JUN-1990 (first entry)
XX Amino acid sequence of human mannose binding protein (hMBP) as encoded by
DE liver cDNA.
XX Human mannose binding protein; liver cDNA.
XX Homo sapiens.
XX WO8901519-A.
PD 23-FEB-1989.
XX 05-AUG-1988; 88WO-US002591.
XX 20-AUG-1987; 87US-00087628.
XX (CHIL-) CHILDRENS HOSPITAL MEDICAL CENT.
XX Ezekowitz RAB;
XX WPI; 1989-068874/09.
XX N-PSDB; AAN91079.
PT New DNA encoding human mannose binding protein or fragment - for treating
PT infections, used for deriving diagnostic antibodies.
XX Disclosure; Fig 2; 34pp; English.
XX The patent claims cDNA encoding 20 or more contiguous amino acids of
CC hMBP, (expression) vectors and cells contg. it, peptides encoded by it,
CC and antibodies (Ab) against these peptides. Peptides encoded by it, opt.
CC fused to a toxin, are used to treat animals infected with bacterial,
CC fungal or viral (including HIV) pathogens. They bind to mannose units on
CC such organisms and cause host defence cells to be attracted to them. The
CC peptides inhibit growth or infection (esp. in HIV, where concn. is 1-500
CC microgram/ml. final serum concn. in human) of the organism. The
CC antibodies are useful in diagnosing susceptibility to infection (by
CC ELISA). (Updated on 25-MAR-2003 to correct PD field.) (Updated on 25-MAR-
XX 2003 to correct PA field.)
SQ Sequence 247 AA;

Alignment Scores:
Pred. No.: 9,65e-98 Length: 247
Score: 1137.00 Matches: 229
Percent Similarity: 92.37% Conservative: 1
Best Local Similarity: 91.97% Mismatches: 15
Query Match: 83.30% Indels: 6
DB: 1 Gaps: 0

US-10-054-536-2 (1-747) x AAP94680 (1-247)

QY 8 TGTTCCTCATCTCCCT 67
DB 3 CysPhe-IleThrProSer-LeuLeuLeuSerMetValAlaAlaSerTySerGluThrV 22
QY 68 TGACCTGTGAGGATGCCCAAAAGACCTGCCCTGCAGTGATTCCTGTAGCTCTCCAGGCA 127
DB 22 aIThrCysGluGlyAlaGlnLysThrCysProAlaValIleAlaCysSerSerProGlyI 42
QY 128 TCACGGCTTCCAGGCAAAAGATGGCGCTGATGGCACCAAGGAGAAAAGGGGGACCGA 187
DB 42 leAsnGlyPheProGlyLysAspGlyArgAspGlyThrLysGlyArgLysGlyThrG 62
QY 188 GCCAAGGGCTCAGAGCTTACAGGGCCCTCCCTGGAAAGTTGGGGCTCCAGGAATCCAG 247
DB 62 lyGlnGlyLeuArgGlyLeuGlnGlyProProGlyLysLeuGlyProProGlyAsnProG 82
QY 248 GGCCTTCTGGGTACACAGGACCAAAAGGGCCAAAAGGAGACCCCTGGAAAAAGTCCGGATG 307
DB 82 lyProSerGlySerProGlyProLysGlyGlnLysGlyAspProGlyLysSerProAspG 102

XX	28-JUL-2003	(first entry)
DT		
XX	Human surfactant protein-D (endothelial form).	
DE		
XX	Human surfactant protein-D; SP-D; atherosclerosis; antiarteriosclerotic;	
KW	antidiabetic; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..21
FT		/label= Leader
FT	Region	22..45
FT		/label= N-terminal
FT	Region	46..129
FT		/label= Collagen
FT	Modified-site	90
FT		/note= "N-glycosylated"
FT	Region	130..157
FT		/label= Neck
FT	Domain	158..282
FT		/label= Carbohydrate_recognition
XX		
PN	WO2003035683-A2.	
XX		
XX	01-MAY-2003.	
PD		
XX		
PP	25-OCT-2002; 2002WO-DK000711.	
XX		
PR	26-OCT-2001; 2001US-0330664P.	
PR	27-DEC-2001; 2001US-0342372P.	
XX		
XX	(HOLM/) HOLMSKOV U.	
PA	(LYST/) LYSTER G L.	
PA	(MADS/) MADSEN J.	
PA	(TORN/) TORNOE I.	
XX		
PI	Holmskov U, Lyster GL, Madsen J, Tornoe I;	
XX		
DR	WPI; 2003-421403/39.	
DR	N-PSDB; ACC57755.	
XX		
PT	New nucleotide sequence encoding surfactant protein-D, useful for	
PT	preparing a composition for preventing or treating diabetes or	
PT	atherosclerosis.	
XX		
PS	Claim 3; Fig 1; 78pp; English.	
XX		
CC	The present sequence is the protein sequence of a novel form of human	
CC	surfactant protein D (SP-D) that is expressed in endothelial cells,	
CC	inhibits oxidation of low density lipoprotein and protects endothelial	
CC	cells from oxidative damage. The endothelial form polypeptide lacks 93	
CC	amino acids compared to the lung form of SP-D and has only 28 Gly-Xaa-Xaa	
CC	repeats as compared to 57 in the lung form of SP-D. The invention	
CC	provides a novel method of diagnosing a person's susceptibility for	
CC	having an increased risk for the development of atherosclerosis by	
CC	measuring the amount of SP-D in plasma and serum samples. An increased	
CC	risk for the development of atherosclerosis in a human is considered at a	
CC	detected serum or plasma SP-D concentration below 1104.6 ng/ml, in	
CC	particular below 743 ng/ml, and especially below 370 ng/ml or 152.7	
CC	ng/ml. SP-D, or a nucleic acid encoding it, can be used to prevent or	
CC	treat an atherosclerosis-related disease such as stroke, kidney failure,	
CC	blindness, leg amputation and myocardial infarction, and to treat obesity	
CC	or diabetes	
XX		
SQ	Sequence 282 AA;	
	Alignment Scores:	
	Pred. No.:	1.39e-35
	Score:	474.00
	Percent Similarity:	52.65%
	Best Local Similarity:	40.28%
		Matches: 114
		Conservative: 35
		Mismatches: 96

Alignment Scores:

Pred. No.: 1.12e-33 Length: 375
 Score: 454.00 Matches: 99
 Percent Similarity: 55.07% Conservative: 26
 Best Local Similarity: 43.61% Mismatches: 82
 Query Match: 33.26% Indels: 20
 DB: 5 Gaps: 5

US-10-054-536-2 (1-747) x AAU76473 (1-375)
 QY 118 TCTCCAGCATCAACGGCTTCCAGGCAAGATGGCGTGATGGCACCAGGAGAAAG 177
 Db AAlaProGlyMetGlnGlySerAlaGlyAlaAargGlyLeuAlaGlyProLysGlyGluArg 168
 QY 178 GGGGAACACAGGCCAA-----GGGCTCAGAGCTTACAGGGCCCCCTCGGAAG 225
 Db GlyValProGlyGluArgGlyValProGlyAsnThrGlyAlaAAlaGlySerAlaGlyAla 188
 QY 226 TTGGGGCTCCAGAAATCCAGGCCCTTCTGGTCACAGGACCAAGGGCCAAAGGA 285
 Db MetGlyProGlnGlySerProGlyAlaAargGlyProGlyLeuLysGlyAspLysGly 208
 QY 286 GACCTCGA---AAAGTCCGGATGCTCATAGTAGCCTG-----GCTGCTCAGAAAGA 336
 Db IleProGlyAspLysGlyAlaLysGlyGlySerGlyLeuProAspValAlaSerLeuArg 228
 QY 337 AAAGCT-----CTGCAACAGAAATGGCAGCTATC 366
 Db GlnGlnValGluAlaLeuGlnGlyGlnValGlnHisLeuGlnAlaAlaPheSerGlnTyr 248
 QY 367 AAAAGTGGGTGACCTTCTCTGGGCAACAAAGTTGGGAACAAGTTCTTCTGACCAAT 426
 Db LysLysValGluLeuPheProAsnGlyGlnSerValGlyGluLysIlePheLysThrAla 268
 QY 427 GGTGAAATATGACCTTGTGAAAAGTGAAGCCCTTGTGTCAAGTCCAGGCTCTGTG 486
 Db GlyPheValLysProPheThrGluAlaGlnLeuLeuCysThrGlnAlaGlyGlyGlnLeu 288
 QY 487 GCCACCCCGAGGAATGCTGCAGAGATGGAGCCATTGAGAAATCTCATC-----AAG 537
 Db AlaSerProArgSerAlaAlaGluAsnAlaAlaLeuGlnGlnLeuValValAlaLysAsn 308
 QY 538 GAGGAAGCTTCTGGGCATCATGTATGAGAAGACAGAGGCGCATGTTGGATCTGACA 597
 Db GluAlaAlaPheLeuSerMetThrAspSerLysThrGluGlyLysPheThrTyrProThr 328
 QY 598 GGAATATAGACTGACATACAACTGGAACGAGGTGACCAACCAATGCTGTTCTGAT 657
 Db GlyLysSerLeuValTyrSerAsnThrPalProGlyGluProAsnAspAspGlyGlySer 348
 QY 658 GAAGATTGTGATTGCTGTAATAATGGCAAAATGGCAGTGAATGACGTCCCTGCTCCACCTCC 717
 Db GluAspCysValGluIlePheThrAsnGlyLysTrpAsnAspArgAlaCysGlyGluLys 368
 QY 718 CATCTGGCCCTCTGTGATGTC 738
 Db ArgLeuValValCysGluPhe 375

RESULT 14

ID ABR44021
 AC ABR44021 standard; protein; 375 AA.
 XX AC ABR44021;
 XX DT 04-AUG-2003 (first entry)
 XX DE Human surfactant protein D (SP-D).
 XX KW Surfactant protein D; SP-D; antiinflammatory; antiasthmatic; rSPD (n/CRD);
 XX KW dermatological; antiallergic; antibacterial; virucide; fungicide; human;
 XX KW antiatherosclerotic; immunosuppressive.
 OS Homo sapiens.

XX FN W02003035679-A2.
 XX PD 01-MAY-2003.
 XX PF 25-OCT-2002; 2002WO-GB004824.
 XX PR 25-OCT-2001; 2001GB-00025638.
 XX PR 26-APR-2002; 2002GB-00009619.
 XX PA (MEDI-) MEDICAL RES COUNCIL.
 XX PI Clark H, Nadesalingam P, Reid KBM, Strong P;
 XX DR WPI; 2003-430400/40.
 XX DR N-PSDB; ACC48169.

Novel recombinant surfactant protein D, termed rSPD n/CRD polypeptide, or its fragment, homolog, variant or derivative, useful for treating inflammatory diseases, asthma, allergy, or atherosclerosis.

Disclosure; Page 126-127; 167pp; English.

The invention relates to a recombinant surfactant protein D (rSPD) (n/CRD) polypeptide (I), or its fragment, homologue, variant or derivative, for use in treatment or prophylaxis of a disease. (I) is useful for treating an individual suffering from a disease or preventing the occurrence of a disease in an individual, where the disease comprises an inflammatory disease, preferably eczema or an inflammatory lung disease (selected from neonatal chronic lung disease, neonatal respiratory distress syndrome (RDS), adult respiratory distress syndrome, chronic obstructive airways disease (COPD), asthma, cystic fibrosis, pulmonary fibrosis, emphysema, interstitial inflammatory lung disease, sarcoidosis, pneumonia, chronic inflammatory lung disease, neonatal chronic inflammatory lung disease) and allergy (see ACC48169 for a detailed description of all the diseases that can be treated using (I)). The present sequence represents a human surfactant protein (SP-D)

SQ Sequence 375 AA;

Alignment Scores:

Pred. No.: 1.12e-33 Length: 375
 Score: 454.00 Matches: 99
 Percent Similarity: 55.07% Conservative: 26
 Best Local Similarity: 43.61% Mismatches: 82
 Query Match: 33.26% Indels: 20
 DB: 5 Gaps: 5

US-10-054-536-2 (1-747) x ABR44021 (1-375)

QY 118 TCTCCAGCATCAACGGCTTCCAGGCAAGATGGCGTGATGGCACCAGGAGAAAG 177
 Db AAlaProGlyMetGlnGlySerAlaGlyAlaAargGlyLeuAlaGlyProLysGlyGluArg 168
 QY 178 GGGGAACACAGGCCAA-----GGGCTCAGAGCTTACAGGGCCCCCTCGGAAG 225
 Db GlyValProGlyGluArgGlyValProGlyAsnThrGlyAlaAlaGlySerAlaGlyAla 188
 QY 226 TTGGGGCTCCAGAAATCCAGGCCCTTCTGGTCACAGGACCAAGGGCCAAAGGA 285
 Db MetGlyProGlnGlySerProGlyAlaAargGlyProGlyLeuLysGlyAspLysGly 208
 QY 286 GACCTCGA---AAAGTCCGGATGCTCATAGTAGCCTG-----GCTGCTCAGAAAGA 336
 Db IleProGlyAspLysGlyAlaLysGlyGlySerGlyLeuProAspValAlaSerLeuArg 228
 QY 337 AAAGCT-----CTGCAACAGAAATGGCAGCTATC 366
 Db GlnGlnValGluAlaLeuGlnGlyGlnValGlnHisLeuGlnAlaAlaPheSerGlnTyr 248
 QY 367 AAAAGTGGGTGACCTTCTCTGGGCAACAAAGTTGGGAACAAGTTCTTCTGACCAAT 426
 Db LysLysValGluLeuPheProAsnGlyGlnSerValGlyGluLysIlePheLysThrAla 268

QY 427 GGTGAATAATAGCCTTGTGAAAGTGAAGCCCTTGTGTCAAGTTCAGGCCCTCTGTG 486
Db GlyPheValLysProPheThrGluAlaGlnLeuLeuCysThrGlnAlaGlyGlyGlnLeu 288
QY 487 GCCACCCCGAGGATGTGCAGAGAATGGAGCCATTGAGATCTCATC-----AAG 537
Db AlaSerProArgSerAlaAlaGluAsnAlaAlaLeuGlnLeuValAlaLysAsn 308
QY 538 GAGGAAGCCTTCTCGGCGATCCTGATGAGAGACAGAGGCCAGTGTGTGATCTGACA 597
Db GluAlaAlaPheLeuSerMetThrAspSerLysThrGluGlyLysPheThrTyrProThr 328
QY 598 GGAATAGACTGACTACACAACTGGAACGAGGTGACCAACATGCTGTTCTGAT 657
Db GlyGluSerLeuValTyrSerAsnTrpAlaProGlyGluProAsnAspGlyGlySer 348
QY 658 GAAGATTGTGATTCTACTGAAAAATGGCCAGTGAATGACGTCCCTGTCTCCACCTCC 717
Db GluAspCysValGluIlePheThrAsnGlyLysTrpAsnAspArgAlaCysGlyGluLys 368
QY 718 CATCTGGCGCTGTGTGATTC 738
Db ArgLeuValValCysGluPhe 375

RESULT 15
AAW18780
ID AAW18780 standard; protein; 351 AA.

AC AAW18780;
XX
DT 11-NOV-1997 (first entry)
XX Bovine conglutinin.
XX Bovine; conglutinin; recombinant; collagen; neck; truncation;
KW sugar chain; recognition; region; antiviral; neutralisation; virus;
KW treatment; disease; influenza A.
XX
OS Bos taurus.
XX WO9707210-A1.
XX
PD 27-FEB-1997.
XX
PF 25-JAN-1996; 96WO-JP000173.
XX
PR 17-AUG-1995; 95JP-00209698.
PR 02-OCT-1995; 95WO-JP002035.
XX
PA (FUSO) FUSO PHARM IND LTD.
XX
PI Wakamiya N;
XX
DR WPI; 1997-165294/15.
XX
PT Recombinant conglutinin with truncated collagen region - has viral
PT neutralising activity and can be used as an antiviral drug.
PS Disclosure; Page 22-23; 48pp; Japanese.

XX The present sequence is bovine conglutinin (bCG), from which a novel
CC recombinant bCG can be prepared. The recombinant bCG comprises a
CC truncated collagen region of formula Gly-Xaa-Gly-Xaa-Xaa (where Xaa
CC is any amino acid), and the neck and sugar chain recognition regions of
CC natural bCG. The recombinant bCG has antiviral activity, i.e. virus
CC neutralising activity, and can be used to treat viral diseases,
XX especially influenza A

XX Sequence 351 AA;

Alignment Scores: 1.32e-32 Length: 351
Pred. No.:

Score: 442.50 Matches: 91
Percent Similarity: 56.77% Conservative: 39
Best Local Similarity: 39.74% Mismatches: 74
Query Match: 32.42% Indels: 25
DB: 2 Gaps: 6

US-10-054-536-2 (1-747) x AAW18780 (1-351)
QY 118 TCTCCAGGCATCAACGGCTTCCAGGCAAAAGATGGCGTGTATGCCACCAAGGAGAAAG 177
Db 126 AlaProGlyIleGlnGlyPheProGlyPro-----SerGlyLeuLysGlyGlyLys 142
QY 178 GGGGAACACAGGCCAA-----GGGCTCAGAGCTTACAGGGCCCCCTGGGAAG 225
Db 143 GlyAlaProGlyGluThrGlyAlaProGlyArgAlaGlyValThrGlyProSerGlyAla 162
QY 226 TTGGGGCTCCAGAAATCCAGGCCCTTCTGGGTACACAGCACCAAGGGCCAAAGGA 285
Db 163 IleGlyProGlnGlyProSerGlyAlaArgGlyProGlyLeuLysGlyAspArgGly 182
QY 286 GACCTCGAAAAGT---CCGGATGGTGATAGTACCTGGCTGCC----- 327
Db 183 AspProGlyGluThrGlyAlaSerGlyGluSerGlyLeuAlaGluValAsnAlaLeuLys 202
QY 328 -----TCAGAAAGAAAAGCTCTGCAACAGAGAAATGGCAGTATC 366
Db 203 GlnArgValThrIleLeuAspGlyHisLeuArgArgPheGlnAsnAlaPheSerGlnTyr 222
QY 367 AAAAAGTGGCTGACCTTCTCTGGGCAACAAAGTTGGGAACAAGTTCTTCTGACCAAT 426
Db 223 LysLysAlaValLeuPheProAspGlyGlnAlaValGlyGluLysIlePheLysThrAla 242
QY 427 GGTGAATAATGACCTTTGAAAAGTGAAGCCCTTGTGTCTCAAGTTCAGGCCCTCTGTG 486
Db 243 GlyAlaValLysSerTyrSerAspAlaGluGlnLeuLysArgGluAlaLysGlyGlnLeu 262
QY 487 GCCACCCCGAGGATGTGCAGAGAAATGGAGCCATTGAGATCTCATCAAG----- 537
Db 263 AlaSerProArgSerAlaGluAsnGluAlaValThrGlnMetValArgAlaGlnGlu 282
QY 538 GAGGAAGCCTTCTCGGCGATCCTGATGAGAGACAGAGGGCAGTTTGTGATCTGACA 597
Db 283 LysAsnAlaTyrLeuSerMetAsnAspIleSerThrGluGlyArgPheThrTyrProThr 302
QY 598 GGAATAGACTGACCTACACAACTGGAACAGAGGTGAACCCCAACAATGCT-----GGT 651
Db 303 GlyGluIleLeuValTyrSerAsnTrpAlaAspGlyGluProAsnAsnSerAspGluGly 322
QY 652 TCTGATGAAGATTGTGTATTGCTACTGAAAAATGGCCAGTGAATGAGTCCCTGCTCC 711
Db 323 GlnProGluAsnCysValGluIlePheProAspGlyLysTrpAsnAspValProCysSer 342
QY 712 ACCTCCCATCTGGCGTCTGTGATTC 738
Db 343 LysGlnLeuLeuValIleCysGluPhe 351

Search completed: June 18, 2004, 21:00:11
Job time : 62.5 secs

B/GNK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 17:16:09 ; Search time 2543 Seconds
(without alignments)
8771.942 Million cell updates/sec

Title: US-10-054-536-2

Perfect score: 747

Sequence: 1 atgtccctgtttccatcact.....tctgtgagttccctatctga 747

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	678	90.8	747	29	AY413286 Homo sapi
2	671.6	89.9	747	29	AY413287 Pan trogl
3	352.2	47.1	735	29	AY413288 Mus muscu
4	335	44.8	828	10	BF537636 602054516

5	303	40.6	367	9	AV6600023	
6	302.2	40.5	657	9	AI255533	
7	301.4	40.3	1162	11	AY325174	
8	301.4	40.3	1162	11	AY325178	
c	9	292.4	39.1	869	9	AI195233
c	10	291	39.0	367	9	AV660367
c	11	290.4	38.9	789	9	AI174038
c	12	286	38.3	1565	11	AK034788
c	13	269	36.0	898	11	BC012245
c	14	268.4	35.9	557	10	BE682267
c	15	268	35.9	788	12	BI147944
c	16	267.6	35.8	784	9	AI173576
c	17	258.4	34.6	777	9	AI194713
c	18	256.6	34.4	664	9	AI876217
c	19	256	34.3	592	9	AI786738
c	20	255.6	34.2	832	14	CB948709
c	21	248.4	33.3	627	14	CF179019
c	22	246.2	33.0	738	12	BI144334
c	23	245.8	32.9	540	13	BX526094
c	24	243.4	32.6	753	12	BI145807
c	25	242.6	32.5	799	9	AI097858
c	26	239.8	32.1	677	9	AI118440
c	27	239.4	32.0	476	10	BF652528
c	28	239	32.0	653	14	CB419261
c	29	236.4	31.6	758	9	AV649316
c	30	234.8	31.4	393	9	AV649531
c	31	231.6	31.0	760	14	CB947613
c	32	228.8	30.6	629	14	CB448619
c	33	227.8	30.5	555	14	CB464046
c	34	225.6	30.2	696	9	AI876428
c	35	218.8	29.3	520	12	BI219712
c	36	216.4	29.0	531	14	CB420319
c	37	213.6	28.6	558	9	AI255561
c	38	211.6	28.3	588	14	CB462601
c	39	210.6	28.2	815	10	BF232803
c	40	208.8	28.0	558	14	CB448348
c	41	207.8	27.8	490	14	CF180387
c	42	206	27.6	700	12	BG084210
c	43	197	26.4	488	9	AA238657
c	44	196.4	26.3	607	9	AI529337
c	45	195.4	26.2	549	9	AW104610

ALIGNMENTS

RESULT 1	AY413286	747 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	AY413286	Homo sapiens MBL2 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	AY413286	genomic survey sequence.			
ACCESSION	AY413286	GI:39769248			
VERSION	AY413286.1				
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	(bases 1 to 747)			
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trics				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2	(bases 1 to 747)			
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				

JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA									
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.									
FEATURES	Location/Qualifiers									
source	1..747									
organism	/organism="Homo sapiens"									
mol_type	/mol_type="genomic DNA"									
db_xref	/db_xref="taxon:9606"									
gene	<1..>747									
	/gene="MBL2"									
	/locus_tag="HCM4840"									
ORIGIN										
Query Match	90.8%;	Score	678;	DB	29;	Length	747;			
Best Local Similarity	90.8%;	0;	Mismatch	180;						
Matches	678;	Conservative	0;	Indels	69;	Indels	0;	Gaps	0;	
Qy	1	ATGTCCTGTTTCCATCACTCCCTCTCTTCTCTGAGTAGTGCGGAGCGCTTACTCA	60							
Db	1	ATGTCCTGTTTCCATCACTCCCTCTCTTCTCTGAGTAGTGCGGAGCGCTTACTCA	60							
Qy	61	GAAACTGTGACCTGTGAGGATGCCAAAAGACCTGCCCTGCAGTGCAITTCCTGTAGTCT	120							
Db	61	GAAACTGTGACCTGTGAGGATGCCAAAAGACCTGCCCTGCAGTGCAITTCCTGTAGTCT	120							
Qy	121	CCAGGCATCAACGGCTTCCAGGCAAGATGGGGTGATGCGCACCAAGGGAGAAAAGGG	180							
Db	121	CCAGGCATCAACGGCTTCCAGGCAAGATGGGGTGATGCGCACCAAGGGAGAAAAGGG	180							
Qy	181	GAACGAGCCAAAGGCTCAGAGGCTTACAGGGCCCCCTGGAAAGTTGGGGCCCTCCAGGA	240							
Db	181	GAACGAGCCAAAGGCTCAGAGGCTTACAGGGCCCCCTGGAAAGTTGGGGCCCTCCAGGA	240							
Qy	241	AATCCAGGGCCTTCTGGGTCAACGAGCACAAAGGGCCAAAAGGAGACCTCGGAAAAAGT	300							
Db	241	AATCCAGGGCCTTCTGGGTCAACGAGCACAAAGGGCCAAAAGGAGACCTCGGAAAAAGT	300							
Qy	301	CCGGATGGTAGTAGCTGGCTGGCTCAGAAAGAAAGCTCTGCAACAGAAATGGCA	360							
Db	301	CCGGNN	360							
Qy	361	CGTATCAAAAAGTGGCTGACCTTCTCTGGGCAACAAAGTTGGGAAACAAGTTCTCTCGT	420							
Db	361	NNNNNNNNNNNGCTGACCTTCTCTGGGCAACAAAGTTGGGAAACAAGTTCTCTCGT	420							
Qy	421	ACCAATGGTGAATAATGACCTTTGAAAAGTGAAGGCCTTGTGTCAAGTTCAGGCC	480							
Db	421	ACCAATGGTGAATAATGACCTTTGAAAAGTGAAGGCCTTGTGTCAAGTTCAGGCC	480							
Qy	481	TCTGTGGCCACCCCGAGGAATGCTGCAGAGAAATGGAGCCATTACAGAAATCTCATCAAGGAG	540							
Db	481	TCTGTGGCCACCCCGAGGAATGCTGCAGAGAAATGGAGCCATTACAGAAATCTCATCAAGGAG	540							
Qy	541	GAAGCCTTCTCTGGGCATCACTGTAGAGAGACAGAGGGCAGTTTGTGGATCTGACAGGA	600							
Db	541	GAAGCCTTCTCTGGGCATCACTGTAGAGAGACAGAGGGCAGTTTGTGGATCTGACAGGA	600							
Qy	601	AATGAGCTGACCTACAAACTGAAACGAGGGTGAACCCCAACAAATGCTGGTCTGATGAA	660							
Db	601	AATGAGCTGACCTACAAACTGAAACGAGGGTGAACCCCAACAAATGCTGGTCTGATGAA	660							
Qy	661	GAITGTGTAITGTCTACTGAAAAATGGCCAGTGAATGACGTCCTCTCTCCACCTCCCAT	720							
Db	661	GAITGTGTAITGTCTACTGAAAAATGGCCAGTGAATGACGTCCTCTCTCCACCTCCCAT	720							
Qy	721	CTGGCCGCTCTGTGAGTTCCCTATCTGA	747							
Db	721	CTGGCCGCTCTGTGAGTTCCCTATCTGA	747							

LOCUS AY413287 747 bp DNA linear GSS 17-DEC-2000
DEFINITION Pan troglodytes MBL2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AV413287
KEYWORDS AV413287.1 GI:39769249
SOURCE GSS.
ORGANISM Pan troglodytes (chimpanzee)
REFERENCE Pan troglodytes
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G.; Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE Pan troglodytes
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G.; Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES
source Location/Qualifiers
 1..747
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>747
 /genes="MBL2"
 /locus_tag="HCM4840"

ORIGIN

Query Match 89.9%; Score 671.6; DB 29; Length 747;
Best Local Similarity 90.2%; Pred.No. 2.le-178;
Matches 674; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 1 ATGTCCTGTGGTTCATCACCCTCCCTCTCTCTCTCTAGTAGTGCGCACGCTTTACTCA 60
Db 1 ATGTCCTGTGGTTCATCACCCTCCCTCTCTCTCTAGTAGTGCGCACGCTTTACTCA 60

Qy 61 GAAACTGTGCCTGTGAGGATGCCAAAAGACTGCCCTGCACTGATTGCCTGTAGCTCT 120
Db 61 GAAACTGTGCCTGTGAGGATGCCAAAAGACTGCCCTGCACTGATTGCCTGTAGCTCT 120

Qy 121 CCAGGCATCAACGGCTCCCAGCAAGATGGCGCTGATGGCCACCAGGGAGAAAAAGGG 180
Db 121 CCAGGCATCAACGGCTCCCAGCAAGATGGCGCTGATGGCCACCAGGGAGAAAAAGGG 180

Qy 181 GAACACAGGCCAAGGCTTCAGAGCTTTACAGGGCCCCCTGGAAAGTTGGGGCTCCAGGA 240
Db 181 GAACACAGGCCAAGGCTTCAGAGCTTTACAGGGCCCCCTGGAAAGTTGGGGCTCCAGGA 240

Qy 241 AATCCAGGGCCTTCTGGGTCACAGGACCAAAGGGCCAAAAGGAGACCTCGGAAAAAGT 300
Db 241 AATCCAGGGCCTTCTGGGTCACAGGACCAAAGGGCCAAAAGGAGACCTCGGAAAAAGT 300

Qy 301 CCGGATGGTGATAGTAGCTGGCTGCCTCAGAAGAAAAGCTCTGCAAACACAGAAATGGCA 360
Db 301 |||||

Qy 361 CGTATCAAAAAGTGGCTGACCTTCTCTGGGCAACAAGTTGGGAACAAGTTCTTCCTG 420
Db 361 NNNNNNNNNNNGGCTGACCTTCTCTGGGCAACAAGTTGGGAACAAGTTCTTCCTG 420

Qy 421 ACCAAATGGTGAATAATAGCTTTGAAAAAGTGAAGGCCTTGTGTGTAAGTTCAGGCC 480
Db 421 |||||

```
Db 421 ACCAATGGTGAATAATGACCTTTGAAAAAGTGAAGGCCCTTGTTGTCTCAAGTTCCAGGCC 480
Qy 481 TCTGTGGCCACCCCGAGGAATGCTGACAGAAATGGAGCCATTTCAGAAATCTCATCAAGGAG 540
Db 481 TCTGTGGCCACCCCGAGGAATGCTGACAGAAATGGAGCCCTTTCGGAATCTCATCAAGGAG 540
Qy 541 GAAGCTCTTCTGGGCATCACTGATGAGAAGACAGAAAGGCGAGTTTCTGGATCTGACAGGA 600
Db 541 GAAGCTCTTCTGGGCATCACTGATGAGAAGACAGAAAGGCGAGTTTCTGGATCTGACAGGA 600
Qy 601 AATAGACTGACCTACACAACTGGACGAGGGTGAAACCCAACTGCTGGTCTCTGATGAA 660
Db 601 AATAGACTGACCTACACAACTGGACGAGGGTGAAACCCAACTGCTGGTCTCTGATGAA 660
Qy 661 GATTGTGATTGCTACTGAAAAATGSCCAGTGAATGAGTCCCTGCTCCACCTCCCAT 720
Db 661 GATTGTGATTGCTACTGAAAAATGSCCAGTGAATGAGTCCCTGCTCCACCTCCCAT 720
Qy 721 CTGGCCGCTGTGAGTTCCTCTCTGA 747
Db 721 CTGGCCGCTGTGAGTTCCTCTCTGA 747

RESULT 3
AY413288
LOCUS 735 bp DNA linear GSS 17-DEC-2003
DEFINITION Mus musculus MBL2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY413288
VERSION AY413288.1 GI:39769250
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 735)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..735
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>735
/gene="MBL2"
/locus_tag="HCM4840"

gene
Query Match 47.1%; Score 352.2; DB 29; Length 735;
Best Local Similarity 66.3%; Pred. No. 4.2e-88;
Matches 479; Conservative 0; Mismatches 237; Indels 6; Gaps 2;

Qy 26 TCCTTCTCCTGATGATGGTGGCAGCGTCTTACTCAGAACTGACCTGTGAGATGCC 85
Db 20 TCCTTCTCCTGATGATGGTGGCAGCGTCTTATGACAGACCTTAACC--GAAGGTGTC 76
Qy 86 AAAGACCTCCCTGTCAGTATTGCCCTGTAGCTCTCCAGGCATCAACGGCTTCCAGGCA 145

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
77 AAAATTCCTGCCCT---GTGGTTACTGTCAGTTCTCCAGGCTGTAATGGCTTCCAGGCA 133
Qy 146 AAGATGGGCGTGTATGGCACCAGAGGAGAGAAAAGGGGAGACAGGCCAAGGGCTCAGAGGCT 205
Db 134 AAGATGGGCGTGTATGGCACCAGAGGAGAGAAAAGGGGAGACAGGTCAGGGCTCAGAGGCT 193
Qy 206 TACAGGGCCCCCTCGAAAAGTTGGGCGCTCCAGGAAATCCAGGGCTTCTGGGTCAACCAG 265
Db 194 TGAAGGCCCTCTCTGAAAAGTAGTACACCTTACAGGACCCCGAGGAATCCGGGTTAAAAG 253
Qy 266 GACCAAGAGGCCCAAAAAGGAGACCCCTGGAAAAGTCCGGATGGTGTATAGTACCTGGCTG 325
Db 254 GAGCAGTGGGACCGAAAAGGAGACCGTGGGGGACAGACGACGNNNNNNNNNNNNNNNN 313
Qy 326 CTTGAAAAGAAAAGCTCTGCAAAAGAAATGGCAGATCAAAAAGTGGCTGACCTTCT 385
Db 314 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 373
Qy 386 CTCTGGGCAACCAAGTTGGGAAACAAGTTCTCTGACCAATGGTGAATAAATGACCTTTG 445
Db 374 CTCTGAGTGAAGAAGTTGGAAGAAGTATTTGTGAGCAGTGTAAAAGATGAGGCTTG 433
Qy 446 AAAAGCTGAAGCCTTGTGTCAAGTTCAGGCTCTTGGCCCTCTTGGCCACCCCGAGGAATGCTG 505
Db 434 ACAGAGTGAAGGCCCTGTGCTCCGAATTCAGGGCTCTGTGGCCACTCCAGGAATGCTG 493
Qy 506 CAGAGATGAGGCCATTTCAGATCTCATCAAGGAGGAAGCCTTCTCTGGGATCAGTATG 565
Db 494 AGGAAAACCTCGGCCATCCAGAAAAGTGGCCAAAGATATTGCTACTTGGGCATCACAGATG 553
Qy 566 AGAAGACAGAAAGGCGCAGTTTGTGGATCTGCAGGAAATAGACTGACTACACAAAATGGA 625
Db 554 TGAGGTTGAAGCAGTCTTTGAGGATCTGCAGGAAACAGAGTGCCTATCTATTTGGA 613
Qy 626 ACAGGGTGAACCCCAACAATGCTGGTTCATGTAAGAATTTGTGTATGCTACTGAAAAATG 685
Db 614 ATGATGGGAGGCCCAACAACACACGGGCGATGGGGAAGACTGTGTGGTGTATCTGGGAAATG 673
Qy 686 GCAGTGGAAATGACGTCCTGCTGCTCCACCTCCATCCCTGAGTTCCTGAGTTCCTATCT 745
Db 674 GCAAGTGGAAACGATGTCCCTGCTCTGACTCTTTTGGCAATCTGTGAATCTCTGACT 733
Qy 746 GA 747
Db 734 GA 735

RESULT 4
BF537636
LOCUS 828 bp mRNA linear EST 11-DEC-2000
DEFINITION 602054516F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4193883 5',
mRNA sequence.
ACCESSION BF537636
VERSION BF537636.1 GI:11625004
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 828)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
```

Plate: L1AM9526 row: o column: 04
 High quality sequence stop: 757.
 Location/Qualifiers
 1. 828
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="PVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4193883"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP SG2"
 /note="Organ: Salivary gland; Vector: pCMV-SPORT6; Site 1:
 Note: Site 2: Sali; Cloned unidirectionally. Primer: Oligo
 dT. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 44.8%; Score 335; DB 10; Length 828;
 Best Local Similarity 71.5%; Pred. No. 3.3e-83;
 Matches 511; Conservative 0; Mismatches 195; Indels 9; Gaps 5;

QY 34 CTGAGTATGGGAGCGCTTACTCAGAACTGTGAGCTGTGAGGATGCCCAAGACC 93
 Db |||||
 QY 3 CTGTGTGTGGTGACAGTGGTTATGCAGAGACCTTAACC--GAAGGTGTTCAAAATTC 59
 Db |||||
 QY 94 TGCCCTGCAGTGATTGCTGTAGCTCTCCAGGCATCAACGGCTCCAGGCAAGATGGG 153
 Db |||||
 QY 60 TGCCCT---GTGGTTACCTGCAGTCTCCAGCCCTGAATGCTTCCAGGCAAGATGGA 116
 Db |||||
 QY 154 CGTGATGCCACCAAGGGAGAAAAGGGGAAACAGGCCAAGGGCTCAGAGGCTTACAGGGC 213
 Db |||||
 QY 117 CGTGACGGTGCACAGGGAGAAAAGGGAGAACAGGTCAAGGGCTCAGAGGCTTCAAGGC 176
 Db |||||
 QY 214 CCCCTGGAAGTTGGGGCTCCAGGAAATCCAGGGCTTCTGGGTACAGGACCAAG 273
 Db |||||
 QY 177 CCTCTGGAAGTAGTAGACCTACAGGACCCCGAGGAATCCGGGGTTAAAGAGGACAGT 236
 Db |||||
 QY 274 GGCCAAAAGAGACCCCTGGAAAAGTCCGATGGTGATAGTGGCTGGCTGCCTCAGAA 333
 Db |||||
 QY 237 GGACGGAAGAGACCGTGGGACAGACAGANTTGATAGGAAATGATTCAGAA 296
 Db |||||
 QY 334 AGAAAAGCTTGCAAAACAGAAATGGACGTATCAAAAAGTGGCTGACCTTCTCTGGGC 393
 Db |||||
 QY 297 ATTGCAGCCCTACGATCAGAGCTGAGAGCCCTGAGAACTGGGTGCTCTTCTCTGAGT 356
 Db |||||
 QY 394 AAAAAGCTTGGAAACAGTCTTCTGACCAATGGTGAATAATGACCTTTGAAAAGTG 453
 Db |||||
 QY 357 GAAAAGTTGAAAAGAG-TATTGTGTAGCAGTGTAAAAAGATGAGCCTTGACAGATG 415
 Db |||||
 QY 454 AAGGCTTGTGTCAAGTCCAGGCTCTGTGGCCACCCAGGAATGCTGCAGAGAAAT 513
 Db |||||
 QY 416 AAGGCCCTGTCTCCGATTTCCAG-CTCTGTGGCCATCCAGGAATGCTGAGGAATC 474
 Db |||||
 QY 514 GGAGCCATTCAGATCTCATCAAGGAGGAAGCCTTCTGGGCATCATCTGATGAGAAGACA 573
 Db |||||
 QY 475 TCGGCCATCCAGAAAGTGGCAAGATATTGCTACTTGGGCATCACAGATGTGAGGGTT 534
 Db |||||
 QY 574 GAAGGCAAGTTGTGATCTCAGAGAAATAGACTGACCTACACAACTGGAACAGGGT 633
 Db |||||
 QY 535 GAAGCAGTTTGTAGGATCTCAGAGAAACAGAGTGGCTATATCTAATTTGAATGATGGG 594
 Db |||||
 QY 634 GAACCCAAATGTGTGTTCTGATGAAGATTGTGTATTGCTACTGAAAAATGGCCAGTGG 693
 Db |||||
 QY 595 GAGCCCAACACACAGGGCGATGGGAGACGTGTGTGTGATCTTGGGAATGGCAAGTGG 654
 Db |||||
 QY 694 AATGACGTCCCT-GCTCCACCTCCCATCTGGCGGTCTGTGAGTTCCTCTATCTGA 747
 Db |||||
 QY 655 AACGATGTCCCTGGCTCTGACTCTTTTGTGCAATCTGTGAAATCTCTGACTGA 709
 Db |||||

RESULT 5

AV660023 367 bp mRNA linear EST 16-JAN-2002
 LOCUS

DEFINITION
 AV660023
 VERSION
 AV660023.1 GI:9881037
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens

REFERENCE
 AUTHORS

Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
 Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
 Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
 Hu,G., Gu,J., Chen,Z. and Han,Z.
 Insight into hepatocellular carcinogenesis at transcriptome level
 by comparing gene expression profiles of hepatocellular carcinoma
 with those of corresponding noncancerous liver
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 21625106
 11752456
 Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chc.sh.cn
 This clone is available at CHGC in Shanghai.
 Location/Qualifiers
 1. 367
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="GLCGDC05"
 /tissue_type="corresponding non cancerous liver tissue"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /clone_lib="GLC"
 /note="vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
 XhoI"

FEATURES
 source

1. 367
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="GLCGDC05"
 /tissue_type="corresponding non cancerous liver tissue"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /clone_lib="GLC"
 /note="vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
 XhoI"

ORIGIN

Query Match 40.6%; Score 303; DB 9; Length 367;
 Best Local Similarity 87.2%; Pred. No. 2.2e-74;
 Matches 367; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

QY 174 AAAGGGGAACAGGCCAAGGGCTCAGAGGCTTACAGGGCCCCCTGAAAAGTTGGGGCC 233
 Db |||||
 QY 1 AAAGGGGAACAGGCCAAGGGCTCAGAGGCTTACAGGGCCCCCTGAAAAGTTGGGGCC 60
 Db |||||
 QY 234 TCCAGGAAATCCAGGGCTTCTGGGTCCAGGACCAAGGGCCAAAGAGAGAGCCCTGG 293
 Db |||||
 QY 61 TCCAGGAAATCCAGGGCTTCTGGGTCCAGGACCAAGGGCCAAAGAGAGAGCCCTGG 120
 Db |||||
 QY 294 AAAAAGTCCGATGGTGATAGTAGCTGGCTGCCTCAGAAAAGAAAAGCTCTGCAACACA 353
 Db |||||
 QY 121 AAAAAGTCCGATGGTGATAGTAGCTGGCTGCCTCAGAAAAGAAAAGCTCTGCAACACA 180
 Db |||||
 QY 354 AATGGCAGTATCAAAAGTGGCTGACCTTCTCTCTGGGCAACAAAGTTGGGACAAGTT 413
 Db |||||
 QY 181 AATGGCAGTATCAAAAGTGGCTGACCTTCTCTCTGGGCAACAAAGTTGGGACAAGTT 240
 Db |||||
 QY 414 CTTCCTGACCAATGGTGAATAATGACCTTTGAAAAGAGTGAAGCCCTTGTGTCAAGTT 473
 Db |||||
 QY 241 CTTCCTGACCAATGGTGAATAATGACCTTTG----- 272
 Db |||||
 QY 474 CCAGGCTCTGTGGCCACCCCGCAAGAAATGTGACAGAAATGGAGCCATTCAGAAATCTCAT 533
 Db |||||
 QY 273 -----ATGCTGCAGAAATGGAGCCATTCAGAAATCTCAT 306
 Db |||||
 QY 534 CAAGGAGGAAGCCTTCTGGGCATCATCTGATGAGAAAGAGAGAGGGCAGTTTGTGGATCT 593
 Db |||||
 QY 307 CAAGGAGGAAGCCTTCTGGGCATCATCTGATGAGAAAGAGAGAGGGCAGTTTGTGGATCT 366
 Db |||||

/translation="MSLFTSLLCLVLTAVYAETLTEGAQSSCPVIACSSPGLNGFPFG
KDHGDAKEGEPEQGLRGLQGGPKGVPAGPKNPGSKGATGPKGRGSEVFDTT
NIDLEIALRSELRAKRWLLSMSENVKVKYFMSVRMPLNRAKALCSELOGTAT
PRNAEENAIQNVADVAFLGTDORTENVEDLTGNRVYTNWNEGRSLRGLPLHG
PAANSVYTPSSSDTENIPTEETAKDGLDPYEQSLIHAATESKASENQHYTLEK
QSVHISVVSVEFLVNLQMFVAVTSSHITGRNRRSLTGLTHEDFPKQADSPSLELTG
WLDRASEPAGCSGLPSGSTRIVDVQVYLDTMTFV"

ORIGIN

```
Query Match      40.3%; Score 301.4; DB 11; Length 1162;
Best Local Similarity 70.6%; Pred. No. 1.3e-73;
Matches 401; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 65 CTGTGACCTGTGAGGATGCCAAAAGACCTGCCCTGCAGTGATTCCTCTGAGTCTCCAG 124
Db 93 CCGAGACCTTAACCGAAGGGGCTCAAGTAGTGCCTGTGATTCCTCTGAGTCTCCGG 152

QY 125 GCATCAACGGCTTCCAGGCAAAAGATGGCGGTGATGGCAACCAAGGAGAAAAGGGGAAC 184
Db 153 GCCTGAACGGCTTCCAGGCAAAAGATGGCAACCAAGGAGAAAAGGGGAAC 212

QY 185 CAGGCCAAGGGCTCAGAGGCTTACAGGGCCCCCTGGAAAGTGTGGGGCTCCAGGAATC 244
Db 213 CGGGTCAAGGCTCAGAGGCTTGCAGGGCCCTCTCTGGAAAAGTAGGACCTGAGGGCCCC 272

QY 245 CAGGGCTTCTGGGTCAACGAGCAAAAGGGCCAAAAGGAGAGACCTCGAAAAGTCCGG 304
Db 273 CAGGGAATCCTGGGTCAAAAGGAGCAACCGGACCAAAAGGAGACCGTGAGAGAGTAG 332

QY 305 ATGGTATAGTAGCTGCTCAGAGAAAGAGTCTGCAAAAGAGTCTGCAAAAGAGTCTGCA 364
Db 333 AATTGATACCTACCAACATTTAGAAATTTGACGCGCTGCGATCGGAGCTGAGAGCTA 392

QY 365 TCAAAAGTGGTCACTTCTCTGGGCAACAGTTGGGACACAGTTCTTCTCTGACCA 424
Db 393 TGAGAAAGTGGTCTCTTCTATAGTGAATGTTGGAAGAAAGTAGTCTATGAGCA 452

QY 425 ATGGTGAATTAATGACCTTTGAAAAGTGAAGGCTTGTGTCTCAAGTTCCAGGCTCTG 484
Db 453 GTGTTAGAGGATGCCCTTAACAGAGGAGGCTGTGTCTCGAATCTCAGGGCACTG 512

QY 485 TGGCCACCCAGGAATGCTGAGAGAAAGTGGAGCAATTCAGAAATCTATCAAGGAGGAAG 544
Db 513 TGGCCACTCCAGGAATGCTGAGGAAAATAGGGCCATCCAGAAATGTGGCCAAAGATGTTG 572

QY 545 CTCTCTGGGATCACTGATCAGAGGACAGAGGCGAGTCTGTGATCTGACAGCAATA 604
Db 573 CTCTCTGGGATCACTGATCAGAGGACAGAGGCGAGTCTGTGATCTGACAGCAATA 632

QY 605 GACTGACCTACCAAACTGGAAACGAGGG 632
Db 633 GAGTGGCGTACACTAACTGGAATGAGGG 660
```

RESULT 8

AY325178 1162 bp mRNA linear HTC 26-JUL-2003
LOCUS Rattus norvegicus Ab2-001 mRNA, complete cds.
DEFINITION AY325178
ACCESSION AY325178
VERSION AY325178.1 GI:33086533
KEYWORDS HTC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1162)

REFERENCE

Xu, C.S., Li, W.Q., Li, Y.C., Ma, H., Wang, L., Wang, S.F., Han, H.P.,
Wang, G.P., Chai, L.Q., Yuan, J.Y., Yang, K.J., Yan, H.M., Chang, C.F.,
Zhao, L.F., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.,
Liver regeneration after PH

JOURNAL

Unpublished

2 (bases 1 to 1162)

AUTHORS

Xu, C.S., Li, W.Q., Li, Y.C., Ma, H., Wang, L., Wang, S.F., Han, H.P.,
Wang, G.P., Chai, L.Q., Yuan, J.Y., Yang, K.J., Yan, H.M., Chang, C.F.,
Zhao, L.F., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.

TITLE

Direct Submision

JOURNAL

Submitted (16-JUN-2003) Henan Bioengineering Key Lab, Henan Normal

University, NO. 148 Jianshe Road, Xinxiang City, Henan 453002, P.R.

China

FEATURES

Location/Qualifiers

1..1162

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

41..1156

/note="Liver regeneration-related protein LRRG053"

/codon_start=1

/product="Ab2-001"

/protein_id="AAP92579.1"

/db_xref="GI:33086534"

/translation="MSLFTSLLCLVLTAVYAETLTEGAQSSCPVIACSSPGLNGFPFG

KDHGDAKEGEPEQGLRGLQGGPKGVPAGPKNPGSKGATGPKGRGSEVFDTT

NIDLEIALRSELRAKRWLLSMSENVKVKYFMSVRMPLNRAKALCSELOGTAT

PRNAEENAIQNVADVAFLGTDORTENVEDLTGNRVYTNWNEGRSLRGLPLHG

PAANSVYTPSSSDTENIPTEETAKDGLDPYEQSLIHAATESKASENQHYTLEK

QSVHISVVSVEFLVNLQMFVAVTSSHITGRNRRSLTGLTHEDFPKQADSPSLELTG

WLDRASEPAGCSGLPSGSTRIVDVQVYLDTMTFV"

ORIGIN

```
Query Match      40.3%; Score 301.4; DB 11; Length 1162;
Best Local Similarity 70.6%; Pred. No. 1.3e-73;
Matches 401; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 65 CTGTGACCTGTGAGGATGCCAAAAGACCTGCCCTGCAGTGATTCCTCTGAGTCTCCAG 124
Db 93 CCGAGACCTTAACCGAAGGGGCTCAAGTAGTGCCTGTGATTCCTCTGAGTCTCCGG 152

QY 125 GCATCAACGGCTTCCAGGCAAAAGATGGCGGTGATGGCAACCAAGGAGAAAAGGGGAAC 184
Db 153 GCCTGAACGGCTTCCAGGCAAAAGATGGCAACCAAGGAGAAAAGGGGAAC 212

QY 185 CAGGCCAAGGGCTCAGAGGCTTACAGGGCCCCCTGGAAAGTGTGGGGCTCCAGGAATC 244
Db 213 CGGGTCAAGGCTCAGAGGCTTGCAGGGCCCTCTCTGGAAAAGTAGGACCTGAGGGCCCC 272

QY 245 CAGGGCTTCTGGGTCAACGAGCAAAAGGGCCAAAAGGAGAGACCTCGAAAAGTCCGG 304
Db 273 CAGGGAATCCTGGGTCAAAAGGAGCAACCGGACCAAAAGGAGACCGTGAGAGAGTAG 332

QY 305 ATGGTATAGTAGCTGCTCAGAGAAAGAGTCTGCAAAAGAGTCTGCAAAAGAGTGGCACGTA 364
Db 333 AATTGATACCTACCAACATTTAGAAATTTGACGCGCTGCGATCGGAGCTGAGAGCTA 392

QY 365 TCAAAAAGTGGTCACTTCTCTGGGCAACAGTTGGGCAACAGTCTTCTCTGACCA 424
Db 393 TGAGAAAAGTGGGTCTCTCTTCTATGATGAAAATGTTGGAAGAAAGTAGTCTATGAGCA 452

QY 425 ATGGTGAATTAATGACCTTTGAAAAGTGAAGGCTTGTGTCTCAAGTTCCAGGCTCTG 484
Db 453 GTGTTAGAGGATGCCCTTAACAGAGGAGGCTGTGTCTCGAATCTCAGGGCACTG 512

QY 485 TGGCCACCCAGGAATGCTGAGAGAAAGTGGAGCAATTCAGAAATCTATCAAGGAGGAAG 544
Db 513 TGGCCACTCCAGGAATGCTGAGGAAAATAGGGCCATCCAGAAATGTGGCCAAAGATGTTG 572

QY 545 CTCTCTGGGATCACTGATCAGAGGACAGAGGCGAGTCTGTGATCTGACAGCAATA 604
Db 573 CTCTCTGGGATCACTGATCAGAGGACAGAGGCGAGTCTGTGATCTGACAGCAATA 632

QY 605 GACTGACCTACCAAACTGGAAACGAGGG 632
Db 633 GAGTGGCGTACACTAACTGGAATGAGGG 660
```

RESULT 9

	Query Match	38.9%;	Score 290.4;	DB 9;	Length 789;
	Best Local Similarity	69.8%;	Pred. No. 1.3e-70;		
	Matches 432;	Conservative 0;	Mismatches 184;	Indels 3;	Gaps 3;
QY	130	AACGGCTTCCAGGCAAAAGATGGCGTGTATGCCACCAAGGGAGAAAAGGGGGAACACAGCG	189		
Db	789	AATGGCTTTCCAGCAAAAATGACCGT-NCGGTGGCCAGGGAGAAAAAGNAGAACCA-GT	732		
QY	190	CAAGGGCTCAGAGGCTTACAGGGCCCCCTCGAAAGTTGGGGGCTCTCAGGAAATCCAGGG	249		
Db	731	CAAGGCTCAANAGGCTTGCAAGCCCTCTCGAAAGTAGGACCTACAGGACCCCCAGGG	672		

671	Db	GATCCGGNGTTAAAGGAGCAGTGGGACCGAAGAGGACCTGGGACAGAGCAAAATT	612
309	QY	TGATAGTAGCCTGGCTGCCTCAGAAAGAAAGCTCTGCAAAACAGAAATGGCCAGTATCAA	368
611	Db	TGATACTAGCGAAATTGATTTCAGAAATTGCAGCCCTACGATCAGAGCTGAGAGCCCTGAG	552
369	QY	AAAGTGGCTGACCTTCTCTCTGGGCAACAAAGTTGGGACACAGTTCTCTCTGACCAATGG	428
551	Db	AAACTGGGGTCTCTTCTCTCTGAGTGAAGAAAGTTGGAAGAAAGTATTTTGTGAGCAGTGT	492
429	QY	TGAAATTAATGACCTTTTGAAAAAGTCAAGGCCCTGTGTGTCAAGTTCCAGGCCCTCTGTGC	488
491	Db	TAAAGATGAGCCTTGCAGAGTGAAGCCCTGTGCTCCGAATTCAGGGCTCTGTGGC	432
489	QY	CACCCCCAGGAATGCTGCAGAGAAATGGAGCCATTCAGAATCTCATCAAGGAGGAAGCCTT	548
431	Db	CACCTCCAGGAATGCTGAGGAAAACTCGGCCATCCAGAAAGTGGCCCAAGATATTGGCTTA	372
549	QY	CCTGGGCATCACTGATGAGAGACAGAGGGCAGTTTGTGGATCTGACAGAGAAATAGACT	608
371	Db	CTTGGGCATCAGATGTGAGGGTTGAAGGCAGTTTTTGAGGATCTCGACAGGAAAACAGAGT	312
609	QY	GACCTACACAACTCGAAACGAGGGTGAACCAACAATGCTGGTCTGATGAAGATTGGT	668
311	Db	GGCCTATACTAAATTGGAATGATGGGAGCCCAACAACACGGGCGGATGGGGAAGACTGTGT	252
669	QY	ATTGCTACTGAAAAATGGCCAGTGAATGAGCTCCCTGCTCCACTCCCACTCTGGCCGT	728
251	Db	GGTGATCTTGGAAATGGCAAGTGGAACGATGTCCTCTGCTCTGACTCTTTTGTGGCAAT	192

QY 729 CTGTGAGTTCCTCTACTGA 747
 ||||| ||| || |||||
 Db 191 CTGTGAATTCCTGACTGA 173

RESULT 12
 AK034788
 LOCUS
 DEFINITION
 Mus musculus 12 days embryo embryonic body between diaphragm region and neck cDNA, RIKEN full-length enriched library, clone:9430039D04
 product:mannose binding lectin, liver (A), full insert sequence.

AK034788
 AK034788.1 GI:26084196
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20493974
 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 695-690 (2001)

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1565)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN, Division of Experimental Animal Research in RIKEN contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1. 1565
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:9430039D04"
 /db_xref="MGI:2398998"
 /db_xref="taxon:10090"
 /clone="9430039D04"
 /tissue_type="embryonic body between diaphragm region and neck"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="12 days embryo"
 120. .840
 /note="mannose binding lectin, liver (A) (MGI:96923, GB|NM_010775, evidence: BLASTN, 99%, match=922) putative"

misc_feature
 38.3%; Score 286; DB 11; Length 1565;
 Best Local Similarity 68.0%; Pred. No. 3 5e-69;
 Matches 416; Conservative 0; Mismatches 190; Indels 6; Gaps 1;

ORIGIN
 Query Match 38.3%; Score 286; DB 11; Length 1565;
 Best Local Similarity 68.0%; Pred. No. 3 5e-69;
 Matches 416; Conservative 0; Mismatches 190; Indels 6; Gaps 1;
 142 GGCAGAGATGGCGTGTATGGCCACCAAGGAGAGAAAGGGGAGAACCGCCAGAGGCTCAGA 201
 Db 229 GGCAGAGATGGGAGAGATGGACCCAAAGGGGAGAGAGGAGGCTCAAGGGCTCAGG 288
 QY 202 GGTTCACAGGGCCCTCCGAAAGTTGGGGCTCCAGGAATCCAGGGCTTCTGGGTCA 261
 Db 289 GGCTTCAGGGCCCTCCAGGAAATTTGGGGCTCCAGGAAGTTTGGAGCCCTCGGAAGT 348
 QY 262 CCAGGACCAAGGGCCAAAGAGAGACCTCGGAAAGTCCGGATGGTATAGTAGCCTG 321
 Db 349 CCAGGACCAAGGGCCAAAGAGAGACCTCGGAAAGTCCGGATGGTATAGTAGCCTG 408
 QY 322 GC-----TGCCTCAGAAAGAAAGCTTCGCAACAGAAATGCCAGTATCAAAAAGTGG 375
 Db 409 GCAATATGGAGGACAGATAGGATCTCTGAATCAAACTGCAGCTAACCAACAAGTTG 468
 QY 376 CTGACCTTCTCTGGGCAACCAAGTTGGGAACAAGTTCTTCTGACCAATGTGAAATA 435
 Db 469 CATGCCCTTCTCAATGGGCAAAAAGTCTGGGAAGAAGTTGTTGTGACCAACCATGAGAAG 528
 QY 436 ATGACCTTTGAAAAGTGAAGGCTTGTGTGTCAGGCTTCAGGCTTCTGTGGCCACCC 495
 Db 529 ATGCCCTTTTCAAAGTGAAGTCTCTGTGCACAGAGCTCAAGGCACTGTGGCTATCCC 588
 QY 496 AGGAATGCTCAGAGAATGGAGCCATTCAAGATCTCATCAAGAGGAGAGCTTCTCTGGGC 555
 Db 589 AGGAATGCTCAGAGAACAAGGCCATTCAAGAAAGTGGCCACAGGCAATTTGCTTAGGC 648
 QY 556 ATCACTGATGAGAAGCAGAGAGGCGAGTTTGTGATCTGCAGGAATAGACTGACCTAC 615
 Db 649 ATCAGGACGAGGCGAGCTGAAGGCGAGTTTATGATGACGAGGAGGAGGCTCACCTAC 708
 QY 616 ACAAACCTGGAACGAGGCTGAACCAACATGTGTTCTCATGAAGATTTGTATTGCTA 675
 Db 709 AGCAACTGGAAGAAAGGATGAGCAATACCATGGCTCTGGGAGAGACTGTGTCATTATA 768
 QY 676 CTGAAAAATGCCAGTGGAAATGAGTCCCTCGCTCCACCTCCCATCTGGCGCTGTGAG 735

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN, Division of Experimental Animal Research in RIKEN contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1. 1565
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:9430039D04"
 /db_xref="MGI:2398998"
 /db_xref="taxon:10090"
 /clone="9430039D04"
 /tissue_type="embryonic body between diaphragm region and neck"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="12 days embryo"
 120. .840
 /note="mannose binding lectin, liver (A) (MGI:96923, GB|NM_010775, evidence: BLASTN, 99%, match=922) putative"

```

Db      769 TTAGTAATGTTTGTGAATGACATTTCTGTCAAGCTTCCTCAAGGCTGTCTCGAG 828
QY      736 TTCCCTATCTGA 747
      ||||| |||||
Db      829 TTCCAGGCTGA 840

RESULT 13
LOCUS   BC012245
DEFINITION Mus musculus, Similar to mammos binding lectin, liver (A), clone
IMAGE:4193682, mRNA.
ACCESSION BC012245
VERSION   BC012245.1 GI:15126625
KEYWORDS  HTC.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg,R.
Direct Submission
Submitted (06-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapba-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunsaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 24 Row: j Column: 9
This clone has the following problem: frame shifted.
FEATURES             source
    source
        1..898
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4193682"
            /tissue_type="salivary gland, 10 week old female mouse"
            /clone_id="NCI CGAP_S62"
            /lab_host="DH10B"
            /note="Vector: pCMV-SPORT6"

ORIGIN
Query Match      36.0%; Score 269; DB 11; Length 898;
Best Local Similarity 67.2%; Pred. No. 1.6e-64;
Matches 416; Conservative 0; Mismatches 190; Indels 13; Gaps 2;

QY      142 GGCAAGATGGCGTGTGGCACCAGAGGAGAAAAGGGGGAACACAGG-----CCAAAGG 194
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      184 GGCAGAGATGGGAGAGATGGACCCAAAGGGAGGAGGAGAACCAAGGTATAGAGTCAAGG 243
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      195 GCTCAGAGGCTTACAGGGCCCTTGGAAAGTTGGGGCCCTCCAGGAATCCAGGCGCTTC 254
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      244 GCTCAGGGGCTTGACGGGCCCTCCAGGGAAATTTGGGGCCCTCCAGGAAGTGTGGAAGCCC 303
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      255 TGGGTACACAGGACCAAGGCGCCAAAAGGAGACCTCGAAAAGTCCGATGGTGTAG 314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      304 TGGNAGTCCAGGACCAAAAGGCCAAAAGGGGGACCATGGAGACATAGACCATTTGAGGA 363
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY      315 TAGCTGGC-----TGCTCAGAAAGAAAGCTCTGCAAAAGAAAATGGCAGCTATCAA 368
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      364 GAAGCTGGCAATATGGAGGCGAGATAAGGATCCTGAAATCAAAGTTCAGCTAACCAA 423
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      369 AAAGTGGCTGACCTTCTCTCTGGGCAAAACAAGTTGGGAACAAGTTCTTCTGACCAATGG 428
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      424 CAAGTTGCATGCTTCTCAATGGGCAAAAAGTCTGGGAAGAAGTTGTTTGTGACCAACA 483
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      429 TGAATAATAGACCTTTGAAAAAGTGAAGCCCTTGTGTCTCAAGTTCCAGGCTCTGTGGC 488
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      484 TGAGAAGATGCCCCCTTTTCCAAAGTGAAGTCTCTGTGACAGAGCTCCAAAGGCACTGTGGC 543
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      489 CACCCCAAGGATGCTGCAGAGATGGAGCCATTGAGAAATCTCATCAAGGAGGAAGCCTT 548
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      544 TATCCCCAGGAATGCTGAAGAGAACAGGCCATTCAAGAAGTGGCCACAGGCAATGSCCTT 603
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      549 CTGGGCATCACTGATGAGAAGACAGAAAGGCGAGTTTCTGGATCTGCAGGAAAATAGACT 608
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      604 CTTAGGCATCAGGACGAGGCGACTGAAGGCGAGTTTCTGACGACAGGGGGAGGCT 663
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      609 GACCTACAAACTGGAAACGAGGGTGAAACCAACAAGTCTGGTTCTGTGATGAAGATTGTGT 668
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      664 CACCTACAGCAACTGGAAAAGGATGAGCCAAATAACCATGGCTCTGGGGAAGACTGTGT 723
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      669 ATTGCTACTGAAAATGCCAGTGGAAATGACGTCCCTGCTCCACCTCCCATCTGGCCGT 728
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      724 CATTAATATTAGTAATGTTTGTGGAATGACATTTCTCTGTCGAAGTTCCTTCAAGGCTGT 783
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      729 CTGTGAGTTCCCTATCTGA 747
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      784 CTGCGAGTTCCAGCCTGA 802
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 14
BE682267
LOCUS
DEFINITION 180230 MARC 4BOV Bos taurus cdna 5', mRNA sequence.
ACCESSION BE682267
VERSION BE682267.1 GI:10067954
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@enail.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGAGC

Plate: 83 row: O column: 15

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

1..557

FEATURES
source

Search completed: June 21, 2004, 19:06:08
Job time : 2549 secs

B/GMK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 18, 2004, 20:56:36 ; Search time 49.5 Seconds
(without alignments)
9522.913 Million cell updates/sec

Title: US-10-054-536-2
Perfect score: 1365
Sequence: 1 atgtccctgttccatcact.....tctgtgagttccctatctga 747

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=slp
-Q=/cgn2_1/USPTO.spool_p/US10054536/runat_18062004_173455_13584/app.query.fasta_1.903
-DB=SPREMBL_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10054536 @CGN 1.86 @runat_18062004_173455_13584 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPREMBL 25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1073	78.6	236	6	Q28518	macaca mula

2	814.5	59.7	245	6	Q28517	macaca mula
3	808	59.2	240	6	Q9XSW3	sus scrofa
4	684	50.1	224	11	Q9Z294	rattus sp.
5	636	46.6	371	11	Q7TMC0	rattus norv
6	551.5	40.4	254	13	Q98TA4	gallus gall
7	540.5	39.6	238	13	O57451	gallus gall
8	454	33.3	375	4	Q86YK9	homo sapien
9	450	33.0	375	4	Q8TCD8	homo sapien
10	428	31.4	246	13	Q919Q7	carassius a
11	422.5	31.0	369	6	Q663A1	bos taurus
12	419.5	30.7	248	6	Q9T06	ovis aries
13	418.5	30.7	256	13	Q919Q9	cyprinus ca
14	411	30.1	378	6	Q9N1X4	sus scrofa
15	407	29.8	248	6	Q95L88	equus cabal
16	407	29.8	248	6	Q9N0G1	equus cabal
17	406.5	29.8	251	13	Q919Q8	brachydanio
18	396	29.0	248	11	Q9CQ11	mus musculus
19	394.5	28.9	248	4	Q81WL2	homo sapien
20	388.5	28.5	248	4	Q81WL1	homo sapien
21	326	23.9	62	4	Q9HCS8	homo sapien
22	304.5	22.3	271	4	Q9BWP8	homo sapien
23	298.5	21.9	272	11	Q9DC75	mus musculus
24	295.5	21.6	271	13	Q7T0T0	xenopus lae
25	295	21.6	742	11	Q8K4Q8	mus musculus
26	295	21.6	742	11	Q8C979	mus musculus
27	289.5	21.2	277	4	Q9Y6Z7	homo sapien
28	289	21.2	742	11	Q8VIF6	mus musculus
29	280.5	20.5	277	11	Q8CF98	mus musculus
30	280.5	20.5	417	4	Q8TCR2	homo sapien
31	278.5	20.4	742	11	Q8C1C5	mus musculus
32	278.5	20.4	742	4	Q9BYH7	homo sapien
33	278.5	20.4	742	4	Q8WZA4	homo sapien
34	275	20.1	268	4	Q7Z6N1	homo sapien
35	248	18.2	116	6	Q9N1X3	sus scrofa
36	239.5	17.5	88	6	Q8MJA5	macaca mula
37	237.5	17.4	165	6	Q9TUC5	macaca mula
38	234	17.1	117	4	Q8TC19	homo sapien
39	219	16.0	222	13	Q90XB2	gallus gall
40	202	14.8	309	11	Q8VH32	mus musculus
41	201.5	14.8	309	11	Q8VH33	mus musculus
42	199	14.6	104	6	Q9XSP4	sus scrofa
43	198.5	14.5	101	6	Q9XSQ2	ovis aries
44	198	14.5	271	5	Q9VET6	drosophila
45	198	14.5	322	5	Q86NZ7	drosophila

ALIGNMENTS

RESULT 1

Q28518	PRELIMINARY;	PRT;	236 AA.
ID	Q28518		
AC	Q28518;		
DT	01-NOV-1996	(TRENBLrel. 01, Created)	
DT	01-NOV-1996	(TRENBLrel. 01, Last sequence update)	
DE	01-OCT-2003	(TRENBLrel. 25, Last annotation update)	
DE		Mannose-binding protein C (Fragment).	
OS	Macaca mulatta (Rhesus macaque).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;		
OC	Cercopithecoidea; Macaca.		
OX	NCBI_TaxID=9544;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RX	MEDLINE=97031450; PubMed=8877375;		
RA	Mogues T., Ora T., Tauber A.I., Sastry K.N.;		
RT	"Characterization of two mannose-binding protein cDNAs from rhesus		
RT	monkey (Macaca mulatta): structure and evolutionary implications."		
RL	Glycobiology 6:543-550(1996).		
DR	EMBL; L43911; AA848071.1; -		
DR	HSSP; P11226; LHUP.		
DR	GO; GO:0005529; F:sugar binding; IEA.		
DR	InterPro; IPR008160; Collagen.		

DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF01391; Collagen; 1.
 DR Pfam: PF00059; lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
 DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
 FT NON_TER 1
 SQ SEQUENCE 236 AA; 24911 MW; 449CBE887C89732E CRC64;

Alignment Scores:

Pred. No.: 4,63e-93 Length: 236
 Score: 1073.00 Matches: 195
 Percent Similarity: 97.22% Conservative: 15
 Best Local Similarity: 90.28% Mismatches: 6
 Query Match: 78.61% Indels: 0
 DB: 6 Gaps: 0

US-10-054-536-2 (1-747) x Q28518 (1-236)

QY 97 CTGAGTGGTCTGCTGAGTCTCCAGGCATCAACGGCTTCCAGGCAAAAGATGGCGT 156
 Db 21 ProGluValIleAlaCysAsnSerProGlyIleAsnGlyPheProGlyLysAspGlyArg 40
 QY 157 GATGACCAAGGAGAAAGGGGACAGCCAGGGCTCAGAGGCTTACAGGCGCC 216
 Db 41 AspGlyThrLysGlyGluLysGlyGluProGlyGlnGlyLeuArgGlyLeuGlnGlyPro 60
 QY 217 CTGGAAGTGGGCGCTCCAGGAATCCAGGGCTTCTGGGTCCAGGACCAAGGGC 276
 Db 61 ProGlyLysLeuGlyProGlyAsnProGlySerSerGlySerProGlyProLysGly 80
 QY 277 CAAAGAGAGACCTCGAAAGTCCGATGTGATAGTCCGCTGCTCGCTCAGAAAGA 336
 Db 81 GlnLysGlyAspProGlyGluSerProAspCysGluSerSerLeuAlaSerGluArg 100
 QY 337 AAAGCTCTGCAACAGAAATGSCAGTATCAAAAAGTGGCTGACCTTCTCTGGGCA 396
 Db 101 LysAlaLeuGlnThrGluMetAlaArgIleLysLysTrpLeuThrPheSerLeuGlyArg 120
 QY 397 CAAGTGGGAACAAGTCTCTGACCAATGGTGAATAATGACCTTTGAAAGTGAAG 456
 Db 121 GlnValGlyAsnLysPhePheLeuThrAsnGlyGluMetMetThrPheAspLysValLys 140
 QY 457 GCCTTGTGTCAAGTTCAGGCTCTGTGGCCACCCAGGAATGTGTCAGAGAATGGA 516
 Db 141 AlaLeuCysAlaArgPheGlnAlaSerValAlaThrProArgAsnAlaAlaGluAsnArg 160
 QY 517 GCCATTGAGATCTCATCAAGGAGAGCCTTCTGGGCATCCTGATGATGATGATGATG 576
 Db 161 AlaIleGlnAsnLeuLysGluGluAlaPheLeuGlyIleThrAspGluAsnThrGlu 180
 QY 577 GGGCAGTTTGGATCTGACAGGAAATAGACTGACCTACACAACTGGAACGAGGGTGA 636
 Db 181 GlyGluPheValAspLeuThrGlyAsnLysLeuThrTyThrAsnTrpAsnAspGlyGlu 200
 QY 637 CCACAACTCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
 Db 201 ProAsnAlaGlySerAsnGluAspCysValLeuLeuLysAsnGlyLysTrpAsn 220
 QY 697 GAGTCCCTGCTCCACCTCCCATCTGGCGCTGTGAGTTCCTATC 744
 Db 221 AspIleProCysSerSerHisLeuAlaLeuCysGluPheProfile 236

RESULT 2

Q28517 PRELIMINARY; PRT; 245 AA.
 AC Q28517;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 OS Mannose-binding protein A precursor (fragment)
 DE Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=970311450; PubMed=8877375;
 RA Moques T., Ota T., Tauber A.I., Sastry K.N.;
 RT "Characterization of two mannose-binding protein cDNAs from rhesus
 monkey (Macaca mulatta): structure and evolutionary implications.";
 RL Glycobiology 6:543-550(1996).
 DR EMBL; L43912; ABA48070.1; -.
 DR HSSP; P19999; LYTT.
 DR GO; GO:0005259; F:sugar binding; IEA.
 DR InterPro: IPR008160; Collagen.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam; PF01391; Collagen; 1.
 DR Pfam; PF00059; lectin_c; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
 DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
 KW Signal.
 FT NON_TER 1
 FT SIGNAL <1 16 POTENTIAL
 FT CHAIN 17 245 MANNOSE-BINDING PROTEIN A.
 SQ SEQUENCE 245 AA; 25693 MW; AFAF0CFEF3DE6146 CRC64;

Alignment Scores:

Pred. No.: 1.62e-68 Length: 245
 Score: 814.50 Matches: 155
 Percent Similarity: 75.92% Conservative: 31
 Best Local Similarity: 63.27% Mismatches: 56
 Query Match: 59.67% Indels: 3
 DB: 6 Gaps: 2

US-10-054-536-2 (1-747) x Q28517 (1-245)

QY 13 CCATCACTCCCT 72
 Db 1 ProSerPheProValLeuLeuLeuSerValValThrAlaSerCysSerGluThrLysAla 20
 QY 73 TGTGAGATGCCCCAAAAGACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132
 Db 21 CysGluAspAlaGlnLysThrCys--SerValIleThrCysGlyIleProValThrAsn 39
 QY 133 GCCTTCCAGCAAGAATGGCGCTGATGGCCAAAGGAGAGAAAGGGGAAAGGGGCA 192
 Db 40 GlyThrProGlyArgAspGlyArgAspGlyProLysGlyGluLysGlyGluProGlyGln 59
 QY 193 GGGCTCAGAGGCTTACAGGCGCCCTTGGAAAGTGGGGCTTCCAGGAAATCCAGGGCCT 252
 Db 60 GlyLeuArgGlyLeuGlnGlyProGlyLysSerGlyProGlyAsnThrGlyAla 79
 QY 253 TCTGGGTCCACAGGACCAAGGGCCAAAAGGAGACCTGGGAAAGTCCGATGGTAT 312
 Db 80 ProGlyIleProGlyProArgGlyGlnLysGlyAspHisGlyAspAsnSerValAlaGlu 99
 QY 313 AGTAGCTGCTGCTCCTCAGAAAGA-----AAAGCTCTGCAACAGAAATGGCAGCATC 366
 Db 100 AlaLysLeuAlaAsnLeuGluArgGlnLeuGlnSerLeuArgSerGluLeuAspHisMet 119
 QY 367 AAAAGTGGCTGACCTTCTCTGGGCAACAAAGTGGGAAACAAGTTCCTTCGACCAAT 426
 Db 120 LysLysLeuGlnAlaPheSerLeuGlyLysMetSerGlyLysLysLeuPheValThrAsn 139
 QY 427 GGTGAATAATGACCTTTGAAAAGTGAAGGCTTGTGTGTCAAGTTCAGGCTCTGTG 486
 Db 140 GlyGluArgMetProPheSerGluValLysAlaLeuCysAlaGlyLeuGlnAlaThrVal 159
 QY 487 GCCACCCCAAGGAATCTCGAGAAATGGAGCAATTCAGAAATCTCATCAAGGAGAGGCC 546
 Db 160 AlaAlaProLysAsnAlaGluGluAsnLysAlaIleGlnAspValAlaLysAspThrAla 179

Tue Jun 22 08:17:17 2004

QY	547	TTCTCTGGGCATCACTGATGAGAACACACAGAGCGGAGTTTGATGATCTCACAGGAANTAGA	606
Db	180	PhelLeuGlyIleThrAspGluAlaThrCluGlyGlnPheMetTyrValThrGlyGlyArg	199
QY	607	CTGACCTTACAAACTGGAACGACGGGTGAACCCACAAATGCTGGTCTTGATGAACATTGT	666
Db	200	LeuThrTyrSerAsnTrpLysLysAspGluProAsnAspHisGlySerGlyGluAspCys	219
QY	667	GTATTGTCTACTGAAATAATGGCCAGCTGGGAATACGCTCCCTCTGCTCCACCTCCCATCTGCC	726
Db	220	ValIleLeuLeuSerAsnGlyLeuTrpAsnAspIleSerCysThrAlaSerTyrIleAla	239
QY	727	GTCTGTGAGTTCCCT	741
Db	240	ValCysGluPhePro	244

RESULT 3	
9XSW3	
D	Q9XSW3 PRELIMINARY; PRT; 240 AA.
C	Q9XSW3;
T	01-NOV-1999 (trEMBLrel_12, Created)
T	01-NOV-1999 (trEMBLrel_12, Last sequence update)
T	01-OCT-2003 (trEMBLrel_25, Last annotation update)
T	Mannose-binding lectin.
S	Sus scrofa (Pig).
S	Eukaryota; Metazoa;
C	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
X	NCBI_TaxID=9823;
N	[1]
N	SEQUENCE FROM N.A.
P	TISSUE=Liver;
C	Aagh A., Young K.; Stahl G.L.;
A	"Isolation, purification and cloning of a porcine mannose-binding
T	lectin.";
T	Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
L	EMBL; AF164576; AAC45377.1; -.
R	HSP; P11226; IHUP.
R	GO; GO:0005529; F:sugar binding; IEA.
R	GO; GO:0007157; P:heterophilic cell adhesion; IEA.
R	InterPro; IPR008161; Clg_helix.
R	InterPro; IPR008160; Collagen.
R	InterPro; IPR001304; Lectin_C.
R	Pfam; PF01391; Collagen; 2.
R	Pfam; PF00059; lectin_c; 1.
R	ProDom; PD000007; Clg_helix; 1.
R	SMART; SM00034; CLECT; 1.
R	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
R	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
R	Collagen; Lectin.
KW	SEQUENCE 240 AA: 25523 MW: 52BD865A21D3D563 CRC64;

Alignment Scores:	
Created. NO.:	6.66e-69
Score:	808.00
Percent Similarity:	74.90%
Best Local Similarity:	65.18%
Query Match:	59.19%
DB:	6
Length:	240
Matches:	161
Conservative:	24
Mismatch:	54
Indels:	8
Gaps:	3

ITS-10-054-536-2 (1-747) x 09XSW3 (1-240)

1	ATGTCCTGTTTCCATCACTCCCTCTCTCTCTCTAGTATGGTCGACGGCTTACTCA	60	Qy
1	MetSerLeuPheProSerLeuHisLeuLeuLeuLeuValMetThrAlaSerHisThr	20	Db
61	GAACCTGTGA C C T G T G A G G A T G C C C A A A A G A C C T G C C C T C A G T G A T T G C C T T A G C T C T	120	Qy
21	GluThrGluAsnCysGluAspIleGlnAsnThrCys---LeuValIleSerCysAspSer	39	Db
121	CCAGCCTCAACGGCTTCCCAAGCAAGATGGGGCTGATGGCCACCAAGGCAGAAAAAGGG	180	Qy
40	ProGluIleAsnGluLeuProGlyIleValAspGlyLeuAspGlyAlaIysGlyGlyGly	59	Db

QY	181	GAACACAGGCGCAAGGGCTCAGAGCTTACACAGGCGCCCTCGAAAGTGTGGGGCTCCACGA	240
Db	60	GlupProGlyGlnGlyLeuileGlyLeuGlnGlyLeuProGlyMetValGlyProGlnGly	79
QY	241	AATCCAGGCGCTTCTGGGTACACAGACCAAGGGCCAAAAGAGAGACCTCGNAAAAGT	300
Db	80	SerProGlyIleProGlyLeuProGlyLeuLysGlyGlnLysGlyAspSerGlyIleAsp	99
QY	301	CCGGATGGTGATAGTAGCTGGCTGCCTCAGAAAGAAAGCTCTGCACAAACAGAAATGGCA	360
Db	100	Pro-----GlyAsnSerLeuAla-----AsnLeuA-gSerGluLeuAsp	112
QY	361	CGTATCAAAAGTGGCTGACCTTCTCTCTGGGCACAAACAGTTGGGAAACAAGTTCTTCTCTG	420
Db	113	AsnIleLysLysTrpLeuIlePheAlaGlnGlyLysGlnValGlyLysLysLeuTyrrhe	132
QY	421	ACCAATGGTGAATATACCTTTGAAAAAGTGAAGGCTGTGTGTCAAGTTCAGGCC	480
Db	133	ThrAsnGlyLysLysMetSerPheAsnGlyValLysAlaLeuCysAlaGlnPheGlnAla	152
QY	481	TCGTGTGCCACCCCGAGGAATGTGCAGAGAATGGAGCCATTCAGAAATCTCATCAAGAG	540
Db	153	SerValAlaThrProThrAsnSerArgGluasnGlnAlaIleGlnGluLeuAlaGlyThr	172
QY	541	GAAGCCCTTCTGGGCATCATGTATGAGAAACAGACAGACGGCAGTTTGTGGATCTGCACGA	600
Db	173	GluAlaPheLeuGlyIleThrAspGluTyThrGluGlyGlnPheValAspLeuThrGly	192
QY	601	AATGAGCTGACCTACACAACTGGAACGAGCGTGAACCCCAACAATGCTGTTTCTCATGAA	660
Db	193	LysArgValArgTyrGlnAsnTrpPAsnAspGlyGluProAsnAsnAlaAspSerAlaGlu	212
QY	661	GATTGTGTATTGCTACTGAAATAATGGCCAGTGGAGTAACGTCCTCGTCTCCACCTCCCAT	720
Db	213	HisCysValGluIleLeuLysAspGlyLysTrpAsnAspIlePheCysSerSerGlnLeu	232
QY	721	CTGGCGCTCTGTGAGTTCCCT	741
Db	233	SerAlaValCysGluPhePro	239
RESULT 4			
Q9Z294		PRELIMINARY;	PRT; 224 AA.
ID	Q9Z294		
AC	DT	01-MAY-1999 (TRENBLrel. 10, Created)	
DC	DT	01-MAY-1999 (TRENBLrel. 10, Last sequence update)	
DT	DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)	
DE	DE	Mannose-binding protein (Fragment).	
OS	OS	Rattus sp.	
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OX	OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.	
RN	RN	NCBI_TaxID=10118;	
RP	RP	[1]	
RC	RC	SEQUENCE FROM N.A.	
RX	RX	TISSUE=Liver;	
RA	RA	MEDLINE=86196130; PubMed=3009480;	
RT	RT	Drickamer K., Dordal M.S., Reynolds L.;	
RT	RT	"Mannose-binding proteins isolated from rat liver contain	
RT	RT	carbohydrate-recognition domains linked to collagenous tails. Complete	
RT	RT	primary structures and homology with pulmonary surfactant	
RT	RT	apoprotein.";	
RL	RL	J. Biol. Chem. 261:6878-6887(1986).	
RN	RN	[2]	
RP	RP	SEQUENCE FROM N.A.	
RC	RC	TISSUE=Liver;	
RA	RA	Drickamer K., Dordal M.S., Reynolds L.;	
RL	RL	Submittal (JUN-1998) to the EMBL/GenBank/DBJ databases.	
DR	DR	EMBL; AF080507; AAC31936.1; --	
DR	DR	HSSP; P19999; 1RTM.	
DR	DR	GO; GO:0005529; F:sugar binding; IEA.	
DR	DR	InterPro; IPR008160; Collagen.	
DR	DR	InterPro; IPR001304; Lectin_C.	
DR	DR	Pfam; PF01391; Collagen; 1.	

US-10-054-536-2 (1-747) x 057451 (1-238)

QY 40 ATGTGGGAGCTCTTACTCAGAACTGTGACCTGTGAGATGCCCAAGACCTGCCCT 99
Db 1 MetMetAlaThrSerLeuLeuThrThrAspLysProGluGluLysMetTyrSerCysPro 20
QY 100 GCAGTATTGCTGTAGCTCTCCAGGCATCAACGGCTTCCAGGCAAGATGGCGCTGAT 159
Db 21 ---IleIleGlnCysSerAlaProAlaValAsnGlyLeuProGlyArgAspGlyArgAsp 39
QY 160 GGCACCAAGGAGAAAGGGGAGCAGGCCAAGGGCTCAGAGCTTACAGGCCCTCCCT 219
Db 40 GlyProLysGlyGluLysGlyAspProGlyGluGlyLeuArgGlyLeuGlnGlyLeuPro 59
QY 220 GGAAGTTGGGCTCCAGGAATCCAGGCCTTCTGGGTCCAGGACCAAGGCCAA 279
Db 60 GlyLysAlaGlyProGlnGlyLeuLysGlyValGlyProGlnGlyGluLysGlyGln 79
QY 280 AAAGGAGACCTCGGAAAGTCCGGATGGTAT-----AGTAGCCTGGCT 324
Db 80 LysGlyGluArgGlyIleValThrAspAspLeuHisArgGlnIleThrAspLeuGlu 99
QY 325 GCTCAGAAAGAAAGCTCTGCAACAGAAATGGCAGCTATCAAAAGTGGCTGACCTTC 384
Db 100 AlaLysIleArg---ValLeuGluAspAspLeuSerArgTyrLysLysAlaLeuSerLeu 118
QY 385 TCTCTGGGCAACAAGTTGGGAACAGTTCTTCTGACCAATGGTGAAATATGACCTTT 444
Db 119 LysAspValValAsnIleGlyLysMetPheValSerThrGlyLysTyrAsnPhe 138
QY 445 GAAAGTGAAGCCCTGTGTCTCAAGTTCCAGGCTCTGTGGCCACCCCGGAAATGCT 504
Db 139 GluLysGlyLysSerLysCysAlaLysAlaGlySerValLeuAlaSerProArgAsnGlu 158
QY 505 GCAGAGATGGAGCAATCAGAACTCATC-----AAGGAGAGCCCTTCTGGGCATC 558
Db 159 AlaGluAsnThrAlaLeuLysAspLeuIleAspProSerSerGlnAlaTyrIleGlyIle 178
QY 559 ACTGATGAGACAGAGGAGGAGTTGGTGGATCTGACAGGAATAGACTACACAC 618
Db 179 SerAspAlaGlnThrGluGlyArgPheMetTyrLeuSerGlyGlyProLeuThrTyrSer 198
QY 619 AACTGGAACGAGGTGAACCCCAACATGCTGCTGATGAAGATGTGTATGTCTACTG 678
Db 199 AsnTyrLysProGlyGluProAsnAsn---HisLysAsnGluAspCysAlaValIleGlu 217
QY 679 AAAATGGCCAGTGAATGACCTCCCTGCTCCACCTCCCATCTG---GCCCTGTGTAG 735
Db 218 AspSerGlyLysTrpAsnAspLeuAspCysSerAsnSerAsnIlePheIleIleCysGlu 237

RESULT 8
Q86YK9 PRELIMINARY; PRT; 375 AA.

AC Q86YK9; TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Surfactant, pulmonary-associated protein D.
GN SFTPD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY216721; AAO22991.1; -
DR GO; GO:0005523; F.sugar binding; IEA.
DR InterPro; IPR008160; Collagen.

DR InterPro; IPR001304; Lectin C.
DR Pfam; PF01391; Collagen; 3.
DR Pfam; PF00059; lectin c; 1.
DR SMART; SM00034; CUECT; 1.
DR PROSITE; PS00615; C TYPE LECTIN 1; 1.
DR PROSITE; PS0041; C TYPE LECTIN 2; 1.
SQ SEQUENCE 375 AA; 37698 MW; 79380764F2B86B67 CRC64;

Alignment Scores: 3e-34 Length: 375
Pred. No.: 454.00 Matches: 99
Score: 55.07% Conservative: 26
Best Local Similarity: 43.61% Mismatches: 82
Query Match: 33.26% Indels: 20
DB: 4 Gaps: 5

US-10-054-536-2 (1-747) x Q86YK9 (1-375)

QY 118 TCTCCAGGCATCAACGGCTTCCAGGCAAGATGGCGTGTGACCAAGGAGGAGAAAG 177
Db 149 AlaProGlyMetGlnGlySerAlaGlyAlaArgGlyLeuAlaGlyProLysGlyGluArg 168
QY 178 GGGGAACCAAGGCCAA-----GGGCTCAGAGGCTTACAGGGCCCCCTGGAAG 225
Db 169 GlyValProGlyGluArgGlyValProGlyAsnAlaGlyAlaAlaGlySerAlaGlyAla 188
QY 226 TTGGGGCTCCAGGAAATCCAGGCCTTCTGGGTCCAGGACCAAGGCCCAAGAAAGGA 285
Db 189 MetGlyProGlnGlySerProGlyAlaArgGlyProGlyLeuLysGlyAspLysGly 208
QY 286 GACCTTGA-----AAAAGTCCGGATGGTGTATAGTAGCTG-----GCTCCCTCAGAAAGA 336
Db 209 IleProGlyAspLysGlyAlaLysGlyGluSerGlyLeuProAspValAlaSerLeuArg 228
QY 337 AAAGCT-----CTGCAACAGAAATGGCAGCTATC 366
Db 229 GlnGlnValGluAlaLeuGlnGlyGlnValGlnHisLeuGlnAlaAlaPheSerGlnTyr 248
QY 367 AAAAAGTGGCTGACCTTCTCTGGGCAACCAAGTTGGGCAACAGTTCTTCTGACCAAT 426
Db 249 LysLysValGluLeuPheProAsnGlyGlnSerValGlyGluLysIlePheLysThrAla 268
QY 427 GGTGAATATGACCTTTGAAAAGTGAAGGCTTGTGTGTCAAGTTCCAGGCCTCTGTG 486
Db 269 GlyPheValLysProPheThrGluAlaGlnLeuLeuCysThrGlnAlaGlyGlnLeu 288
QY 487 GCCACCCCAAGATGCTGCAGAGATGGAGCATTCAGAACTCATC-----AAG 537
Db 289 AlaSerProArgSerAlaAlaGluAsnAlaAlaLeuGlnGlnLeuValValAlaLysAsn 308
QY 538 GAGGAAGCTTCTCCGGCATCACTGATGAGAAGACAGAGGCGAGTTGTGGATCTGACA 597
Db 309 GluAlaAlaPheLeuSerMetThrAspSerLysThrGluGlyLysPheThrTyrProThr 328
QY 598 GGAATATAGACTGACCTACAACTGGAACCGAGGCTGAACCCCAACAACTGCTGCTGTGAT 657
Db 329 GlyGluSerLeuValTyrSerAsnTrpAlaProGlyGluProAsnAspAspGlyGlySer 348
QY 658 GAAGATTGTGTATCTACTGAAAATGGCCAGTGAATGAGCTCCCTGCTCCACCTCC 717
Db 349 GluAspCysValGluIlePheThrAsnGlyLysTrpAsnAspArgAlaCysGlyGluLys 368
QY 718 CATCTGGCGCTGTGTGAGTTC 738
Db 369 ArgLeuValValCysGluPhe 375

RESULT 9
Q8TCD8 PRELIMINARY; PRT; 375 AA.

ID Q8TCD8; TREMBLrel. 21, Created)
AC Q8TCD8; TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

Q919Q7	PRELIMINARY;	PRT;	246 AA.
ID	Q919Q7		
AC	Q919Q7;		
DT	01-OCT-2000	(TEMBLrel. 15, Created)	
DT	01-OCT-2000	(TEMBLrel. 15, Last sequence update)	
DT	01-OCT-2003	(TEMBLrel. 25, Last annotation update)	
DE	Mannose binding-like lectin precursor (fragment).		
GN	MBL		
OS	Carassius auratus (Goldfish).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		
OC	Cyprinidae; Carassius.		
OX	NCBI_TaxID=7957;		

RP SEQUENCE FROM N.A.
RC
RD
RE TISSUE=Liver;
RF
RX MEDLINE=20456727; PubMed=11003389;
RA Vitved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjoldt K.;
RV "the homologue of mannose-binding lectin in the carp family Cyprinidae
RW is expressed at high level in spleen, and the deduced primary
RX structure predicts affinity for galactose.";
RY
RZ

[illegible]

```

Db 72 ProGlyThrAlaGlyLeuLysGlyGluArgGlyProSerGlyProGlnGlySerProGly 91
QY 295 AAAAGTCCGGATGCTGATAGTAGCTGGCTGCTCAGAAAGAAAGCTCTGCAACAGAA 354
Db 92 SerGluSerValLeuGluSer-----LeuLysSerGluLeuGlnLeuLysAlaLys 109
QY 355 ATGGCAGCTATCAAAAGTGGCTGACCTTCTCTGGGCAACAAAGTGGGAACAAGTTC 414
Db 110 IleAlaThrPheGluLysValSerSerValCysHisPheArgLysValGlyGlnLysTyr 129
QY 415 TTCCTGACCAATGCTGAATAATAGCTTTGAAAGTGAAGCCCTTGTGTCTCAAGTTC 474
Db 130 TyrIleThrAspGlyValValGlyAsnPheAsnGlyLeuLysSerCysMetGluPhe 149
QY 475 CAGGCTCTGTGGCCACCCAGGAATGCTGCAGAGAATGGAGCCATTGAGATCTCATC 534
Db 150 GlyGlyThrMetValSerProArgThrSerAlaGluAsnGlnAlaLeuLysLeuVal 169
QY 535 -----AAGGAGGAGCCTTCTCTGGGCATCCTGAGGAGGAGGAGGAGGAGGAG 576
Db 170 ValSerSerGlyLeuGlySerLysLysProTyrIleGlyValThrAspArgLysThrGlu 189
QY 577 GGCAGATTGCTGCTCAGAGGAATAGCTGCTACAACTGGAACGAGGCTGAA 636
Db 190 GlyGlnPheValAspThrGluGlyLysGlnLeuThrPheThrAsnTrpGlyProGlyGln 209
QY 637 CCCAACATGCTGCTGATGAGATTGTGATTGCTACTGAAAAATGGCCAGTGAAT 696
Db 210 ProAspTyrLysGlyLeuGlnAspCysGlyValIleGluAspThrGlyLeuTrpAsp 229
QY 697 GACGTCCCTGCTCCACTCCCATCTGCGCGTGTGTGAGTTCCTTCATC 744
Db 230 AspGlyGlyCysGlyAspIleArgProIleMetCysGluIleAspIle 245

```

RESULT 11

```

Q863A1 PRELIMINARY; PRT; 369 AA.
AC Q863A1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surfactant protein D precursor.
GN SP-D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Gjerstorff M., Hansen S., Madsen J., Bendixen C., Holmskov U.;
RT "Bovine surfactant protein D: Genomic characterization, chromosomal
RT localization and tissue distribution."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ548848; CAD69922.1; JOINED.
DR EMBL; AJ548849; CAD69922.1; JOINED.
DR EMBL; AJ548850; CAD69922.1; JOINED.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01391; Collagen_3.
DR SMART; SM00039; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
KW SIGNAL.
FT SIGNAL. 1 20 Potential.
FT CHAIN 21 67 surfactant protein D.
FT CHAIN 245 369 surfactant protein D.
SQ SEQUENCE 369 AA; 37405 MW; 4A74B7593508A5ED CRC64;

```

Alignment Scores:

Pred. No.: 2.92e-31 Length: 369

```

Score: 422.50 Matches: 88
Percent Similarity: 52.68% Conservative: 30
Best Local Similarity: 39.29% Mismatches: 89
Query Match: 30.95% Indels: 17
DB: 6 Gaps: 4

```

US-10-054-536-2 (1-747) x Q863A1 (1-369)

```

QY 118 TCTCAGGATCAACGGCTTCCAGGCAAGATGGCGTGATGGCAACCAAGGAGAGAAAG 177
Db 146 AlaProGlyIleGlnGlySerProGlyProAlaGlyLeuLysGlyGluArgGlyAlaPro 165
QY 178 GGGGAACACAGC---CAAGGGCTCAGAGCTTACAGGGCCCTCCCTGGAAAGTTGGGCT 234
Db 166 GlyGluProGlyAlaProGlyArgAlaGlyAlaProGlyProAlaGlyAlaIleGlyPro 185
QY 235 CCAGAAATCCAGGCTTCTGGGTCCAGGACCAAGGCAAGGCAAGGACCAAGGACCTGGA 294
Db 186 GlnGlyProSerGlyAlaArgGlyProGlyLeuLysGlyAspArgGlyThrProGly 205
QY 295 ---AAAATCCGGATGGTAGTAGCTGGCT----- 324
Db 206 GluArgGlyAlaLysGlyGluSerGlyLeuAlaGluValAsnAlaLeuArgGlnArgVal 225
QY 325 -----GCCTCAGAAAGAAAGCTCTGCAACAGAAATGGCAGCTATCAAAAGTGG 375
Db 226 GlyIleLeuGluGlyGlnLeuGlnArgLeuGlnAsnAlaPheSerGlnTyrLysAla 245
QY 376 CTGACCTTCTCTCTGGGCAACAAAGTTGGGAACAAGTTCTTCTCACCACCAATGGTGAATA 435
Db 246 MetLeuPheProAsnGlyArgSerValGlyGluLysIlePheLysThrGluGlySerGlu 265
QY 436 ATGACCTTTGAAAAGTGAAGCCCTTGTGTGTCAAGTTCAGGCTCTGTGGCCACCCCT 495
Db 266 LysThrPheGlnAspAlaGlnGlnIleCysThrGlnAlaGlyGlyGlnLeuProSerPro 285
QY 496 AGGAATGCTCAGAGAATGGAGCCATTGAGATCTCATC-----AAGGAGGAAGCC 546
Db 286 ArgSerAlaAlaGluAsnGlnAlaLeuThrGlnLeuAlaThrAlaGlnAsnLysAlaAla 305
QY 547 TTCCTGGGCTCACTGATGAGACAGAGGCGAGTTGTGGATCTGACAGGAAATAGA 606
Db 306 PheLeuSerMetSerAspThrArgLysGluGlyThrPheIleTyrProThrGlyGluPro 325
QY 607 CTGACCTACAAACTGGAACGAGGGTGAACCCCAACAAATGCTGTCTGTGATGAAGATTGT 666
Db 326 LeuValTyrSerAsnTrpAlaProGlnGluProAsnAsnAspGlyGlySerGluAsnCys 345
QY 667 GTATTGCTACTGAAAATGGCCAGTGAATGAGTCCCTCCCTGCTCCACCTCCCATCTGCC 726
Db 346 ValGluIlePheProAsnGlyLysTrpAsnAspLysValCysGlyGluGlnArgLeuVal 365
QY 727 GTCTGTGAGTTC 738
Db 366 IleCysGluPhe 369

```

RESULT 12

```

Q9TT06 PRELIMINARY; PRT; 248 AA.
ID Q9TT06
AC Q9TT06;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pulmonary surfactant protein A (Pulmonary surfactant-associated
DE protein A).
DE SPAS OR SP-A.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OC NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.

```

493 CCCAGGAATGCTGCAGAGAATGAGCCATTCAAGTCTCATCAAGGAGGAA-----544
 |||||
 165 ProArgSerProGluGluAsnGluAlaIleThrSerileValLysHisAsnThrTyr 184
 |||||
 544 GCCTTCCTGGGCATCCTACTGATGAGACACAGAAGAGCGCAGTTTGTTGGATCTGACAGGAAT 603
 |||||
 185 AlaTyrLeuGlyLeuAlaGluGlyProThrAlaGlyAspPheTyrTyrLeuAspGlyAla 204
 |||||
 604 AGACTGCTACACAAAACCTGGAACGAGGTGAACCCCAACAATGCTGCTGTTCTGATGNAGAT 663
 :||:
 205 ProValAsnTyrThrAsnTrpTyrProGlyGluProArgGlyArgGlyLys---GluLys 223
 |||||
 664 TGTGATTGCTACTGAAAAATGCCAGTGAATGACGTCCTCCCTGCTCCACCTCCCCTCTG 723
 |||||
 224 CysValGluIleTyrThrAspGlyGlnTrpAsnAspLysAsnCysLeuGlnTyrArgLeu 243
 |||||
 724 GCGCTGCTGTGAGTTC 738
 |||||
 244 AlaIleCysGluPhe 248
 |||||

RESULT 13
 Q91909 PRELIMINARY; PRT; 256 AA.
 ID Q91909 AC Q91909
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Mannose binding-like lectin precursor.
 OS MBL.
 GN Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=liver;
 RX MEDLINE=20456722; PubMed=11003389;
 RA Vitved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjoldt K.;
 RT "the homologue of mannose-binding lectin in the carp family Cyprinidae
 RT is expressed at high level in spleen, and the deduced primary
 RT structure predicts affinity for galactose.";
 RL Immunogenetics 51:955-964(2000).
 DR EMBL; AF227737; AAF63468.1; -.
 DR HSSP; P35247; 1B08.
 DR GO; GO:0005529; F:sugar binding; IEA.
 DR GO; GO:0007157; Phetotrophic cell adhesion; IEA.
 DR InterPro; IPR008160; Collagen.
 DR InterPro; IPR001304; Lectin.C.
 DR Pfam; PF01391; Collagen; 2.
 DR pfam; PF00059; lectin.c; 1.
 DR SMART; SM00034; CLECT_1.
 DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 KW Lectin; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 256 MANNOSE BINDING-LIKE LECTIN.
 FT VARIANT 235 235 S->T.
 SQ SEQUENCE 256 AA; 26934 MW; D01929D1167730D CRC64;

Alignment Scores:	
Pred. No.:	6.4e-31
Score:	418.50
Percent Similarity:	54.05%
Best Local Similarity:	36.68%
Query Match:	30.66%
DB:	13
Length:	256
Matches:	95
Conservative:	45
Mismatches:	98
Indels:	21
Gaps:	7

US-10-054-536-2 (1-747) x Q91909 (1-256)
 QY 1 ATGTCCTGTTTCCA----TACTCCCTCTCTCTGATGTTGGCAGCGCTTCTAC 57
 |||||
 Dh 1 MetAlaLeuPheLysLeuPheLeuGlyThrLeuLeuLeuGlnPheAlaLeuGlnLeu 200

QY	58	TCAGAAACGTGTGACCTGTGAGGATGCCCAAAAGACCTGCGCTGCAGTGATTGCCTGT---	114
Db	21	LeuAsp-----GlyAlaGluProGlnAsnLeuAsnCysProAlaTyrGlyGlyValPro	38
QY	115	AGCTCTCCAGGCAATCAACGGCTTCCAGGCAAGATGGCGTGAAT-----	159
Db	39	GlyThrProGlyHisAsnGlyLeuProGlyArgAspGlyArgAspGlyLysAspGlyAla	58
QY	160	---GGCACCAACGGAGAAAGGGGNAACAGGCCAAGGGCTCAGAGGCTTACAGGGCCCC	216
Db	59	IleGlyProLysGlyGluLysGlyGluSerGlyVal-----SerValGlnGlyPro	75
QY	217	CCTGAAAAGTTGGGGCCTCCAGAAATCCAGGGCCTTCTGGGTACACGAGGACCAAGGGC	276
Db	76	ProGlyLysAlaGlyProGlyProAlaGlyGluLysGlyGluArgGlyProThrGly	95
QY	277	CAAAAGAGAGCCCTGGGAAAAGTCCGGATGGTGATAGTACGTGGCTGCCTCAGAAAGA	336
Db	96	SerGlnGlySerProGlySerGluSerValLeuGluSer-----LeuLysSerGluIle	113
QY	337	AAAGCTCTGCAAAACAGAAATCGCAGTATCAAAAAGTGGCTGACCTTCTCTGGGGCAA	396
Db	114	GlnGlnLeuLysAlaLysIleAlaThrPheGluLysValAlaSerValGlyHisPheArg	133
QY	397	CAAGTTGGGAACAAGTTCTTCTGACCAATGGTGAATATATGACCTTTGAAAAGTGAAG	456
Db	134	GlnValGlyGlnLysTyrTyrIleThrAspGlyValValGlyThrPheAspGlnGlyLeu	153
QY	457	GCCTTGTTGTCAAGTTCAGGCCTCTGTGGCCACCCCGAATGCTGCAGAGAATGGA	516
Db	154	LysPheCysLysAspPheGlyGlyThrMetValPheProArgThrSerAlaGluAsnGln	173
QY	517	GCCATTGAGAATCTCATC-----AAGGAGGAAGCCTTCTGGGGCATC	558
Db	174	AlaLeuLeuLysLeuValValSerSerGlyLeuSerSerLysLysProTyrIleGlyVal	193
QY	559	ACTGATGAGAAGACAGAGGGCGATTTGTGGATCTGACAGGAAATACAGCTACCTACAC	618
Db	194	ThrAspArgGluThrGluGlyArgPheValAsnThrGluGlyLysGlnLeuThrPheThr	213
QY	619	AACTGGNAACAGGGTGNAACCCACAATGCTGGTTCTGATGAAGATTTGTATTGCTACTG	678
Db	214	AsnTrpGlyProGlyGlnProAspAspTyr-LysGlyLeuGlnAspCysGlyValIleGlu	233
QY	679	AAAAATGGCCAGTGGGAATGACGTCCTGCTCCACCTCCCATCTGGCCGCTCTGTGAG	735
Db	234	AspSerGlyLeuTrpAspAspGlySerCysGlyAspIleArgProIleMetCysGlu	252

RESULT 14

ID	Q9N1X4	PRELIMINARY;	PRT;	378 AA.
Q9N1X4	Q9N1X4			
AC	Q9N1X4;			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	lung surfactant protein D precursor.			
DE	SPRPD.			
GN	GN			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20109098; PubMed=10640760;			
RA	van Eijk M., Haagsman H.P., Skinner T., Archibald A.L., Reid K.B.M.,			
RA	Lawson P.R.;			
RT	"Porcine Lung Surfactant Protein D (Sp-D): cDNA cloning, chromosomal			
RT	localisation and tissue distribution."			
RL	J. Immunol. 164:1442-1450(2000).			
DR	EMBL; AF132496; AAF22145.2; -.			
DR	HSSP; P35247; 1B06.			
DR	GO; GO:0005529; F:sugar binding; IEA.			

```
Db 239 GlnArgLeuGlnLysAlaPheSerGlnTyrLysLysValGluLeuPheProAsnGlyArg 258
QY 397 CAAGTTGGGAACAAGTTCTCTCGACCAATGGTGAATAATACACCTTTGAAAAGTGAAG 456
Db 259 GlyValGlyLysIlePheLysThrGlyGlyPheGluLysThrPheGlnAspAlaGln 278
QY 457 GCCTTGTGTGCAAGTTCACAGGCTCTGTGGCCACCCAGGAATGCTGCAGAGAATGGA 516
Db 279 GlnValCysThrGlnAlaGlyGlnMetAlaSerProArgSerGluThrGluAsnGlu 298
QY 517 GCATTTCAGNATCTCATC-----AAGGAGGAGCCCTTCTCGGCATCACTGATGAG 567
Db 299 AlaLeuSerGlnLeuValThrAlaGlnAsnLysAlaAlaPheLeuSerMetThrAspIle 318
QY 568 AAGACAGAGGCGCAGTTGTGGATCTGACAGGAATAAGACTGACCTACACAACTGGAAC 627
Db 319 LysThrGluGlyAsnPheThrTyrProThrGlyGluProLeuValTyrAlaAsnTyrAla 338
QY 628 GAGGGTGAACCCCAACATGCTGTTCTGAT-----GAAGATTGTGTATTGCTACTG 678
Db 339 ProGlyGluProAsnAsnGlyGlySerSerGlyAlaGluAsnCysValGluIlePhe 358
QY 679 AAAAAAGCCAGTGAATGAGTCCCTGCTCCACCTCCCATCTGGCCGCTCTGTGATTC 738
Db 359 ProAsnGlyLysTyrPheAsnAspLysAlaCysGlyGluLeuArgLeuValIleCysGluPhe 378

RESULT 15
Q95L88 PRELIMINARY; PRT; 248 AA.
AC Q95L88;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pulmonary surfactant-associated protein A.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N. A.
RA Weber B.I.L., Hospes R., Gortner L.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF400580; AAL07690.1; -.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR00160; Collagen.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF00059; lectin.c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C.TYPE LECTIN 1; 1.
DR PROSITE; PS50041; C.TYPE LECTIN 2; 1.
SQ SEQUENCE 248 AA; 26047 MW; B71133E005C9A501 CRC64;
```

```
Alignment Scores:
Pred. No.: 7.84e-30 Length: 248
Score: 407.00 Matches: 97
Percent Similarity: 51.74% Conservativity: 37
Best Local Similarity: 37.45% Mismatches: 87
Query Match: 29.82% Indels: 38
DB: 6 Gaps: 12
```

US-10-054-536-2 (1-747) x Q95L88 (1-248)

```
QY 28 CTTCTCTGAGTATGGTGGCAGCGTCTTACTCAGAACTGTGACCTGTGAGATGCCCA 87
Db 6 LeuThrLeuThrLeuLeuLeuAlaValSerGlyThr----LysCys---AspValLys 23
QY 88 AAGACCTGCCCTGCAGTATGCTGTGATGCTCT-----CCAGGCATCAACGGC 135
Db 24 GluPheCys-----AlaAlaCysSerGlyValProGlyIleProGlySerProGly 40
QY 136 TTCCAGGCAAGATGGGCGTGTGGCACCAGAGGAGAAAGGGGGAACCCAGGCCAAGG 195
```

```
Db 41 LeuProGlyArgAspGlyValLysGlyAspProGlyProGlyProGlyProGlyProGly 59
QY 196 CTCAGAGGCTTACAGGCCCCCTGAAAGTTGGGCGCTCCAGGAATCCA-----246
Db 60 -----IleGlyProProGly-----GlyMetProGlySerProGlyHisAsp 73
QY 247 -----GGGCTTCTGGGTCAACAGGACCAAGGGGCAAAAGAGGACCTCGA---294
Db 74 GlyLeuIleGlyProProGlyProGlyGluArgGlyAspLysGlyGluProGlyGlu 93
QY 295 AAAAGTCCCGATGTGATAGTAGCTGCTCTCAGAAAGAAAGAAAGCTTGCAGAAACAGAA 354
Db 94 ArgGlyProProGlyProProAlaTyrProAspGluGlu-----LeuGlnThrThr 110
QY 355 ATGGCAGCTATCAAAAAGTGGCTGACCTCTCTCTGGGCAAA-----396
Db 111 LeuHisAspIleArgHisGlnIleLeuGlnLeuMetGlyAlaLeuSerLeuGlnGlySer 130
QY 397 -----CAAGTTGGGAACAAGTTCTTCTGACCAATGGTGAAATAATGACCTTTGAAAAA 450
Db 131 MetLeuAlaValGlyGluLysValPheSerThrAsnGlyGlnValValAspPheAspAla 150
QY 451 GTCAAGGCTTGTGTGTCAAGTTCCAGGCTCTGTGGCCACCCAGGAAATGCTGCAGAG 510
Db 151 IleArgGluSerCysAlaArgAlaGlyGlyArgIleAlaValProLysSerLeuGluGlu 170
QY 511 AATGGAGCCATTCAGAAATCTCATCAAGGAGAA-----GCCTTCTGGGCATCACT 561
Db 171 AsnAlaAlaIleAlaSerLeuValThrLysHisAsnThrTyrAlaTyrLeuGlyLeuGlu 190
QY 562 GATGAGAAGACAGAGGCGAGTTTGTGGATCTCAGAGGAATAGACTGACCTACACAAAC 621
Db 191 GluGlyProThrAlaGlyAspPheTyrTyrLeuAspGlyAlaProValAsnTyrThrAsn 210
QY 622 TGAACACGAGGTTGAACCCCAATGTGTTCTCTGATGAAGATTGTGTATTGCTACTGAAA 681
Db 211 TptTyrProGlyGluProArgGlyArgGlyLys---GluLysCysValGluMetTyrThr 229
QY 682 AATGGCCAGTGAATGACGTCCCTCTGCTCCACCTCCCATCTGGCCGCTCTGAGTTC 738
Db 230 AspGlyGlnTrpAsnAspArgSerCysLeuGlnTyrArgLeuAlaIleCysGluPhe 248
```

Search completed: June 18, 2004, 21:02:45
Job time : 54.5 secs

QY 61 GAACTGTGACCTGTGAGATGCCCAAAAGACCTGCCCTGAGTGAATGCTGCTGCTCT 120
 DB 21 GluThrValThrCysGluAspAlaGlnIysThrCysProAlaValIleAlaCysSer 40
 QY 121 CAGGATCAACGGCTTCCAGCAAGATGGCGTGTGAGCCACCAAGGAGGAGGAGG 180
 DB 41 ProGlyIleAsnGlyPheProGlyIysAspGlyIysAspGlyIysAspGlyIys 60
 QY 181 GAACAGGCGCAAGGCTCAGAGCTTACAGGGCCCCCTGGAAAGTTGGGGCTCCAG 240
 DB 61 GluProGlyGlnGlyLeuArgGlyLeuGlnGlyProGlyIysLeuGlyProGly 80
 QY 241 AATCAGGCGCTTCTGGGTACACAGCACCAAGAGGCGCAAAAGGAGCACCTGG 300
 DB 81 AsnProGlyProSerGlySerProGlyProIysGlyGlnIysGlyAspProGly 100
 QY 301 CGGATGCTGATAGTACCTGCTGCTCAGAAAGAAAGCTCTGCAAAACAGAAATGG 360
 DB 101 ProAspGlyAspSerSerLeuAlaIaSerGluArgIysAlaLeuGlnThrGlu 120
 QY 361 CGTATCAAAAAGTGGCTGACCTTCTCTGGGCAAAACAAAGTTGGGAAACAGTT 420
 DB 121 ArgIleIysIysTrpLeuThrPheSerLeuGlyIysGlnValGlyAsnIysPhe 140
 QY 421 ACCAATGGTGAATAATGACCTTTGAAAAGTGAAGGCTTGTGTCAAGTTCCAGGC 480
 DB 141 ThrAsnGlyGluIleMetThrPheGluIysValIysAlaLeuCysValIysPhe 160
 QY 481 TCTGTGGCCACCCAGGATGCTGCAGAAATGGAGCCCTTGTGTCAAGTTCCAGGC 540
 DB 161 SerValAlaThrProArgAsnAlaIaGluAsnGlyAlaIleGlnAsnLeuIle 180
 QY 541 GAAGCCTTCTGGGCAATCACTGATGAGACACAGAGGAGGAGTTGGATCTGACAG 600
 DB 181 GluAlaPheLeuGlyIleThrAspGluIysThrGluGlyGlnPheValAspLeu 200
 QY 601 AATAGACTGACCTACAAACTGGAACGAGGTGAACCAACATGCTGCTGATGAA 660
 DB 201 AsnArgLeuThrThrAsnTrpAsnGluGlyIleProAsnAsnAlaGlySerAsp 220
 QY 661 GATTGTGTTGCTGCTGAAATGGCCAGTGAATGACGTCCTGCTCCACCTCCCAT 720
 DB 221 AspCysValLeuLeuLeuIysAsnGlyGlnTrpAsnAspValProCysSerThr 240
 QY 721 CTGGCCCTGCTGAGTTCCTTCATC 744
 DB 241 LeuAlaValCysGluPheProIle 248

RESULT 2
 US-09-011-735-1
 ; Sequence 1, Application US/09011735B
 ; Patent No. 6110708
 ; GENERAL INFORMATION:
 ; APPLICANT: Wakamiya, No. 6110708utaka
 ; TITLE OF INVENTION: Recombinant Conglutinin and Producing Method Thereof
 ; FILE REFERENCE: 19036/34548
 ; CURRENT APPLICATION NUMBER: US/09/011,735B
 ; EARLIER FILING DATE: 1998-05-22
 ; EARLIER FILING DATE: 1998-05-22
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 351
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-09-011-735-1
 Alignment Scores:
 Pred. No.: 1.15e-36 Length: 351
 Score: 442.50 Matches: 91
 Percent Similarity: 56.77% Conservative: 39
 Best Local Similarity: 39.74% Mismatches: 74

Query Match: 32.42% Indels: 25
 DB: 3 Gaps: 6
 US-10-054-536-2 (1-747) x US-09-011-735-1 (1-351)
 QY 118 TCTCAGGATCAACGGCTTCCAGGCAAGATGGCGTGTGAGCCACCAAGGAGGAGG 177
 DB 126 AlaProGlyIleGlnGlyPheProGlyPro-----SerGlyLeuIysGlyGlu 142
 QY 178 GGGGACACGAGGCAAA-----GGGCTCAGAGGCTTACAGGGCCCCCTGGAAAG 225
 DB 143 GlyAlaProGlyGluThrGlyAlaProGlyArgAlaGlyValThrGlyProSerGly 162
 QY 226 TTGGGCGCTCAGGAAATCCAGGCGCTTCTGGGTACACAGGACCAAAAGGCGCAAAAGGA 285
 DB 163 IleGlyProGlnGlyProSerGlyAlaArgGlyProProGlyLeuIysGlyAspArg 182
 QY 286 GACCTTGGAAAAAGT---CCGATGCTGATAGTAGCTGGCTGCC----- 327
 DB 183 AspProGlyGluThrGlyAlaSerGlyGluSerGlyLeuAlaGluValAsnAlaLeu 202
 QY 328 -----TCAGAAAGAAAGCTCTGCAAAACAGAAATGGCACCTATC 366
 DB 203 GlnArgValThrIleLeuAspGlyHisLeuArgArgPheGlnAsnAlaPheSerGln 222
 QY 367 AAAAAGTGGCTGACCTTCTCTGGGCAAAACAAAGTTGGGAAACAGTTCTTCTGACCAAT 426
 DB 223 LysIysAlaValLeuPheProAspGlyGlnAlaValGlyGluIysIlePheLysThr 242
 QY 427 GGTGAAATATACCTTTGAAAAAGTGAAGCCCTTGTGTCAAGTTCCAGGCGCTG 486
 DB 243 GlyAlaValIysSerIysAspAlaGluGlnLeuCysArgGluAlaIysGlyGlnLeu 262
 QY 487 GCCACCCCGAGGATGCTGCAGAGATGGAGCCATTCAGAAATCTCATCAAG----- 537
 DB 263 AlaSerProArgSerSerAlaGluAsnGluAlaValThrGlnMetValArgAlaGln 282
 QY 538 GAGGAAGCCTTCTGGGCACTGATGAGAAAGACAGAGGCGAGTTGTGGATCTGACA 597
 DB 283 LysAsnAlaIysLeuSerMetAsnAspIleSerThrGluGlyArgPheThrTyProThr 302
 QY 598 GGAATAGACTGACCTACAAACTGGAACGAGGTGAACCAACCAACATGCT-----GGT 651
 DB 303 GlyGluIleLeuValTySerAsnTrpAlaAspGlyGluProAsnAsnSerAspGlu 322
 QY 652 TCTGATGAAGATCTGTATTCTACTGAAAAATGCCAGTGAATGAGTCCCTGCTCC 711
 DB 323 GlnProGluAsnCysValGluIlePheProAspGlyIysTrpAsnAspValProCys 342
 QY 712 ACCTCCCATCTGGCGCTGTGTGAGTTC 738
 DB 343 LysGlnLeuLeuValIleCysGluPhe 351
 RESULT 3
 US-09-029-156-1
 ; Sequence 1, Application US/09029156
 ; Patent No. 6365342
 ; GENERAL INFORMATION:
 ; APPLICANT: WAKAMIYA, No. 6365342utaka
 ; TITLE OF INVENTION: RECOMBINANT CONGLUTININ AND PRODUCING
 ; TITLE OF INVENTION: METHOD THEREOF
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 South Wacker Drive/6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: PatentIn Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/09/029,156

/ FILING DATE:

/ CLASSIFICATION: 435

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: PCT/JP96/00173

/ FILING DATE:

/ APPLICATION NUMBER: PCT/JP95/02035

/ FILING DATE: 02-OCT-1995

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: JPA - 209698

/ FILING DATE: 17-AUG-1995

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Gass, David A.

/ REGISTRATION NUMBER: 38,153

/ REFERENCE/DOCKET NUMBER: 19036/34546

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (312) 474-6300

/ TELEFAX: (312) 474-0448

/ INFORMATION FOR SEQ ID NO: 1:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 351 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: not relevant

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

/ US-09-029-156-1

Alignment Scores:

Pred. No.:	1,15e-36	Length:	351
Score:	442.50	Matches:	91
Percent Similarity:	56.77%	Conservative:	39
Best Local Similarity:	39.74%	Mismatches:	74
Query Match:	32.42%	Indels:	25
DB:	4	Gaps:	6

US-10-054-536-2 (1-747) x US-09-029-156-1 (1-351)

QY 118 TCTCAGGATCAACGGCTTCCAGCAAGATGGCGGTGATGGACCAAGGAGAAAG 177

Db 126 AlaProGlyIleGlnGlyPheProGlyPro-----SerGlyLeuLysGlyGluLys 142

QY 178 GGGGAACCAAGGCAAA-----GGGCTCAGAGGCTTACAGGGCCCCCTCGAAAG 225

Db 143 GlyAlaProGlyGluThrGlyAlaProGlyArgAlaGlyValThrGlyProSerGlyAla 162

QY 226 TTGGGGCTCCAGGAATCCAGGCTTCTGGTCCAGGACCAAGGGCCAAAGGA 285

Db 163 IleGlyProGlnGlyProSerGlyAlaArgGlyProGlyLeuLysGlyAspArgGly 182

QY 286 GACCTCGGAAAAAGT---CCGGATGGTATAGTACCTGGCTGCC----- 327

Db 183 AspProGlyGluThrGlyAlaSerGlyGluSerGlyLeuAlaGluValAsnAlaLeuLys 202

QY 328 -----TCGAAAGAAAAGCTCTCGAAACAGAAATGGCAGCTATC 366

Db 203 GlnArgValThrIleLeuAspGlyHisLeuArgArgPheGlnAsnAlaPheSerGlnTyr 222

QY 367 AAAAGTGCTGACCTTCTCTGGGCAACAAAGTTGGGAAACAAGTTCTTCCTGACCAAT 426

Db 223 LysLysAlaValLeuPheProaspGlyGlnAlaValGlyGluLysIlePheLysThrAla 242

QY 427 GTTGAATAATGACCTTTGAAAAGTGAAGGCTTGTGTCAAGTTCAGGCTCTGTG 486

Db 243 GlyAlaValLysSerTyrSerAspAlaGluGlnLeuLysCysArgGluAlaLysGlyGlnLeu 262

QY 487 GCCACCCCGGAAATGCTCAGAGAAATGGAGCATCAAGTCTCATCAAG----- 537

Db 263 AlaSerProArgSerSerAlaGluAsnGluAlaValThrGlnMetValArgAlaGlnGlu 282

QY 538 GAGGAAGCTTCTCGGGCATCTACTGATGAGACAGAGGCGAGTTTGTGGATCTGACA 597

Db	283	LysAsnAlaTyrIleuSerMetAsnAspIleSerThrGluGlyArgPheThrTyrProThr	302
QY	598	GGAAATAGACTGACCTACACAACTGGAACGAGGGTGAACCCCAACATGCT-----GGT	651
Db	303	GlyGluIleLeuValTyrSerAsnTrpAlaAspGlyGluProAsnAsnSerAspGluGly	322
QY	652	TCTGATGAAGATTGTGTATTGTCTACTGAAAAATGGCCAGTGGGAATGACGTCCTCCCTGCTCC	711
Db	323	GlnProGluAsnCysValGluIlePheProAspGlyLysIleTrpAsnAspValProCysSer	342
QY	712	ACCTCCCATCTGGCCGCTCTGTGAGTTC 738	
Db	343	LysGlnLeuLeuValIleCysGluPhe 351	

RESULT 4

5514582-16

/ Patent No. 5514582

/ APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.

/ TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID

/ IMMUNOGLOBULINS

/ NUMBER OF SEQUENCES: 43

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/185,670

/ FILING DATE: 21-JAN-1994

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: 986,931

/ FILING DATE: 08-DEC-1992

/ APPLICATION NUMBER: 808,122

/ FILING DATE: 16-DEC-1991

/ APPLICATION NUMBER: 440,625

/ FILING DATE: 22-NOV-1989

/ APPLICATION NUMBER: 315,015

/ FILING DATE: 23-FEB-1989

/ SEQ ID NO: 16:

/ LENGTH: 108

5514582-16

Alignment Scores:

Pred. No.:	2,35e-28	Length:	108
Score:	359.00	Matches:	66
Percent Similarity:	74.07%	Conservative:	14
Best Local Similarity:	61.11%	Mismatches:	28
Query Match:	26.30%	Indels:	0
DB:	6	Gaps:	0

US-10-054-536-2 (1-747) x 5514582-16 (1-108)

QY 409 AAGTTCTTCTGACCAATGGTGAATAATGACCTTTGAAAAAGTGAAGGCTTGTGTGTC 468

Db 1 LysPhePheValThrAsnHisGluArgMetProPheSerLysValLysAlaLeuCysSer 20

QY 469 AAGTTCAGGCTCTGTGGCCACCCCGAGGATGCTGCAGAGAAATGGAGCCATTCAGAT 528

Db 21 GluLeuArgGlyThrValAlaIleProLysAsnAlaGluLysAlaIleGlnGlu 40

QY 529 CTCATCAAGAGGAGAGCTTCTCGGSCATCACTGATGAGAACAGAGGGGAGTTGTG 588

Db 41 ValAlaLysThrSerAlaPheLeuGlyIleThrAspGluValThrGluGlyGlnPheMet 60

QY 589 GATCTCAGAGGAATAGACTGACCTACACAACTCGAAGAGGGGTGACCCCAACATGCT 648

Db 61 TyrValThrGlyGlyArgLeuThrTyrSerAsnTrpLysLysAspGluProAsnAspHis 80

QY 649 GGTTCGTGATGAAGATTGTGTATTGCTACTGAAAAATGCCAGTGAATGACGTCCTCCCTGC 708

Db 81 GlySerGlyGluAspCysValThrIleValAspAsnGlyLeuTrpAsnAspIleSerCys 100

QY 709 TCCACCTCCCATCTGGCGCTCTGT 732

Db 101 GlnAlaSerHisThrAlaValCys 108

RESULT 5

5514582-17

<hr/>					
; Patent No. 5514582					
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.					
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID					
; IMMUNOGLOBULINS					
; NUMBER OF SEQUENCES: 43					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/185,670					
; FILING DATE: 21-JAN-1994					
; PRIOR APPLICATION DATA:					
; APPLICATION NUMBER: 986,931					
; FILING DATE: 08-DEC-1992					
; APPLICATION NUMBER: 808,122					
; FILING DATE: 16-DEC-1991					
; APPLICATION NUMBER: 440,625					
; FILING DATE: 22-NOV-1989					
; APPLICATION NUMBER: 315,015					
; FILING DATE: 23-FEB-1989					
; SEQ ID NO:17:					
; LENGTH: 107					
5514582-17					
Alignment Scores:					
Pred. No.:	3,49e-27	Length:	107		
Score:	347.50	Matches:	66		
Percent Similarity:	73.15%	Conservative:	13		
Best Local Similarity:	61.11%	Mismatches:	28		
Query Match:	25.46%	Indels:	1		
DB:	6	Gaps:	1		
US-10-054-536-2 (1-747) x 5514582-17 (1-107)					
Qy	409	AAGTCTTCGTGACCAATGGTGAATAATGACCTTTGAAAAAGTGAAAGCGCTTGTGTCTC	468		
Dd	1	LysPhePheValThrAsnHisGluArgMetProPheSerLysValLeuAlaLeuCysSer	20		
Qy	469	AAGTCCAGGCCTCTGTGCCACCACCCAGCAATGCTGCAGAGAATGGAGCCATTCAAGAT	528		
Dd	21	GluLeuArgGlyThrValAlaIleProArgAsnAlaGluGluAsnLysAlaIleGlnGlu	40		
Qy	529	CTCATCAAGGAGGAGCGCTTCTGGCATCTGATGAGAACAGAGGCGGCGTGTCTG	588		
Dd	41	ValAlaLysThrSerAlaPheLeuGlyIleThrAspGluValThrGluGlyGlnPheMet	60		
Qy	589	GATCTCACAGGAATAGACTGACCTACACAACACTGACGAGGTGAACCAACAAATGCT	648		
Dd	61	TyrValThrGlyArgLeuThrTyrrSerAsnTrpLysLysAspGluProAsnAspHis	80		
Qy	649	GTTTCTGATGAGATGTTGTATGCTACTGAAAATGGCCAGTGAATGACGTCCCTCC	708		
Dd	81	Gly---HisGluAspCysValThrIleValAspAsnGlyLeuTrpAsnAspIleSerCys	99		
Qy	709	TCCACTCCCATCTGGCCCTGTGT	732		
Dd	100	GlnAlaSerHisThrAlaValCys	107		
RESULT 6					
US-09-011-735-6					
; Sequence 6, Application US/09011735B					
; Patent No. 6110708					
; GENERAL INFORMATION:					
; APPLICANT: Wakamiya, No. 6110708utaka					
; TITLE OF INVENTION: Recombinant Conglutinin and Producing Method Thereof					
; FILE REFERENCE: 19036/34548					
; CURRENT APPLICATION NUMBER: US/09/011,735B					
; CURRENT FILING DATE: 1998-05-22					
; EARLIER APPLICATION NUMBER: JP 7-209698					
; EARLIER FILING DATE: 1995-08-17					
; NUMBER OF SEQ ID NOS: 6					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 6					
; LENGTH: 161					
; TYPE: PRT					
; ORGANISM: Bovine					
; FEATURE:					
; NAME/KEY: VARIANT					
; LOCATION: (2)					
; OTHER INFORMATION: Xaa is a protein-constituting amino acid					
; FEATURE:					
; NAME/KEY: VARIANT					
; LOCATION: (3)					
; OTHER INFORMATION: Xaa is a protein-constituting amino acid					
; FEATURE:					
; NAME/KEY: VARIANT					
; LOCATION: (5)					
; OTHER INFORMATION: Xaa is a protein-constituting amino acid					
; FEATURE:					
; NAME/KEY: VARIANT					
; LOCATION: (6)					
; OTHER INFORMATION: Xaa is a protein-constituting amino acid					
US-09-011-735-6					
Alignment Scores:					
Pred. No.:	6,86e-21	Length:	161		
Score:	286.50	Matches:	54		
Percent Similarity:	61.15%	Conservative:	31		
Best Local Similarity:	38.85%	Mismatches:	49		
Query Match:	20.99%	Indels:	5		
DB:	3	Gaps:	2		
US-10-054-536-2 (1-747) x US-09-011-735-6 (1-161)					
Qy	337	AAGCTCTCGAACAACAGAAATGGCACGTATCAAAAAGTGGCTGACCTTCTCTGGGCAA	396		
Dd	23	ArgArgPheGlnAsnAlaPheSerGlnTyrlsLysAlaValLeuPheProaspGlyGln	42		
Qy	397	CAAAGTGGGAAACAAGTTCTTCTCGACCAATGGTGAATAATGACCTTTGAAAAAGTGAAG	456		
Dd	43	AlaValGlyGluLysIlePheLysThrAlaGlyAlaValLysSerTyrrSerAspAlaGlu	62		
Qy	457	GCCTTGTGTGTAAGTTCACAGGCTCTGTGGCCACCCACGAGGAATGCTGCAGAGATGA	516		
Dd	63	GlnLeuCysArggluAlalysGlyGlnLeuAlaSerProArgSerSerAlaGluAsnGlu	82		
Qy	517	GCCATTCAAGATCTCATCAAG-----GAGGAAGCCTTCTGGCATCACTCATGAG	567		
Dd	83	AlaValThrGlnMetValArgAlaGlnGluLysAsnAlaTyrlsSerMetAsnAspIle	102		
Qy	568	AAGACAGAAGGCGAGTTTCTGGATCTGACAGAAATAGACTACCTACACAACTGGAGAC	627		
Dd	103	SerThrGluGlyArgPheThrTyrrProThrGlyGluLeuValTyrrSerAsnTrpAla	122		
Qy	628	GAGGTGAACCCAACTGCT-----GGTTCTGATGAGATGTGTGCTACTGCTGAA	681		
Dd	123	AspGlyGluProAsnAsnSerAspGlyGlnProGluAsnCysValGluIlePhePro	142		
Qy	682	AATGGCCAGTGAATGACGTCCCCTCTCCCATCTCCCATCTCCCGCTCTGTGAGTTC	738		
Dd	143	AspGlyLysTrpAsnAspValProCysSerLysGlnLeuValIleCysGluPhe	161		
RESULT 7					
5514582-20					
; Patent No. 5514582					
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.					
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID					
; IMMUNOGLOBULINS					
; NUMBER OF SEQUENCES: 43					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/185,670					
; FILING DATE: 21-JAN-1994					
; PRIOR APPLICATION DATA:					
; APPLICATION NUMBER: 986,931					
; FILING DATE: 08-DEC-1992					
; APPLICATION NUMBER: 808,122					

```
; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO: 20:
; LENGTH: 111
5514582-20

Alignment Scores:
Pred. No.: 2,08e-13 Length: 111
Score: 212.50 Matches: 42
Percent Similarity: 56.25% Conservative: 21
Best Local Similarity: 37.50% Mismatches: 44
Query Match: 15.57% Indels: 5
DB: 6 Gaps: 3

US-10-054-536-2 (1-747) x 5514582-20 (1-111)
QY 409 AAGTTCTTCTGACCAATGGTGAATAATGACCTTTGAAAAAGTGAAGGCCCTTGTGTGTC 468
Db 1 LysIlePheSerThrAsnGlyGlnSerValAsnPheAspAlaIleArgGluValCysAla 20
QY 469 AAGTTCAGGCTCTGTGGCCACCCCGAGGAATGCTGCAGAGAAAT---GGAGCCATTGAG 525
Db 21 ArgAlaGlyGlyArgIleAlaValProArgSerLeuGluGluAsnGluAlaIleAla 40
QY 526 AATCTCATCAAGGAGGAA-----GCTTCCTGGGCATCACTGATGAGAAGACAGAA 576
Db 41 SerIleValLysGluArgAsnThrTyrAlaTyrLeuGlyLeuAlaGluGlyProThrAla 60
QY 577 GGGCAGTTTGTGGATCTGACAGGAATAGACTACACAACTGGAAACGAGGGTGAA 636
Db 61 GlyAspPheTyrTyrLeuAspGlyAspProValAsnTyrThrAsnTyrTyrProGlyGlu 80
QY 637 CCCAACATGCTGCTTCTGATGAAGATTCTGATTCTACTGAAAAATGCCAGTGGAA 696
Db 81 ProArgGlyGlnGlyArg---GluLysCysValGluMetTyrThrAspGlyLysTrpAsn 99
QY 697 GACGTCCTGCTGCCACCTCCCATCCCATCGGCCGTCTGT 732
Db 100 AspLysAsnCysLeuGlnTyrArgLeuValIleCys 111

RESULT 8
US-09-602-877A-99
; Sequence 99, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-602-877A-99

Alignment Scores:
Pred. No.: 1.93e-12 Length: 197
Score: 204.00 Matches: 54
Percent Similarity: 46.63% Conservative: 29
Best Local Similarity: 30.34% Mismatches: 65
Query Match: 14.95% Indels: 30
DB: 4 Gaps: 5

US-10-054-536-2 (1-747) x US-09-602-877A-99 (1-197)
QY 271 AAGGGCCAAAAGGAGACCCCTGGAAAGTCCGGATGGTATAGTAGCCTG----- 321
||| :|||
||| :|||
```

```
Db 28 LysAlaArgLysHisSerLysArgValArgAspLysAspGlyAspLeuLysThrGln 47
QY 322 -----GCTGCCTCAGAAAGAAAGCTCTGCAACAGAA 354
Db 48 IleGluLysLeuTrpThrGluValAsnAlaLeuLysGluIleGlnAlaLeuGlnThrVal 67
QY 355 ATGGCAGCTATCAAAAAGTGGCTGACCTTCTCTGGGCGAAACAAGTTGGGAACAAGTTC 414
Db 68 CysLeuArg-----GlyThrLysValHisLysLysCys 78
QY 415 TTCCTGACCAATGTGAAATATGACCTTTGAAAAAGTGAAGCCCTGTGTCTCAAGTTC 474
Db 79 TyrLeuAlaSerGluGlyLeuLysHisPheHisGluAlaAsnGluAspCysLysSerLys 98
QY 475 CAGGCTCTGTGGCCACCCCGAGGAATGCTGCAGAGAATGGAGCCATTCAGAAATCTCATC 534
Db 99 GlyGlyIleLeuValIleProArgAsnSerAspGluIleAsnAlaLeuGlnAspTyrGly 118
QY 535 AAG-----GAGGAAGCCTTCTCTGGGCATCACTGATGAGAAGACAGAA 576
Db 119 LysArgSerLeuProGlyValAsnAspPheTrpLeuGlyIleAsnAspMetValThrGlu 138
QY 577 GGGCAGTTTGTGGATCTGACAGGAATAGACTACACAACTGGAAACGAGGGTGAA 636
Db 139 GlyLysPheValAspValAsnGlyIleAlaIleSerPheLeuAsnTrpAspArgAlaGln 158
QY 637 CCCAACATGCTGCTTCTGATGAAGATTCTGATTGCTACTGAAAAAT-----GGCCAG 690
Db 159 Pro---AsnGlyGlyLysArgGluAsnCysValLeuPheSerGlnSerAlaGlnGlyLys 177
QY 691 TGAATGACGTCCTCCCTGCTCCACCTCCCATCTGGCGCTCTGTGAGTTCCTCATC 744
Db 178 TrpSerAspGluAlaCysArgSerSerLysArgTyrIleCysGluPheThrIle 195

RESULT 9
US-08-365-103B-4
; Sequence 4, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G
; APPLICANT: Nunez, Raphael D.
; APPLICANT: Yodoi, Jungi
; TITLE OF INVENTION: DNA Sequences for Soluble Proms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Uirf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 287 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
US-08-365-103B-4

Alignment Scores:
Pred. No.: 4,04e-12 Length: 287
Score: 201.50 Matches: 51
Percent Similarity: 46.59% Conservative: 31
Best Local Similarity: 28.98% Mismatches: 47
Query Match: 14.76% Indels: 47
DB: 1 Gaps: 7

US-10-054-536-2 (1-747) x US-08-365-103B-4 (1-287)
QY 316 AGCTGCTGCTCCTCAGAAAGAAAGCT-----CTGCAACAGAA 354
D 108 SerLeuGlyLeuAsnGluLysArgThrAlaSerAspSerLeuGluLysLeuGlnGlu 127
QY 355 ATGCACGATC----- 366
D 128 ValAlaLysLeuTrpIleGluLeuSerLysGlyThrAlaCysAsnIleCysPro 147
QY 367 AAAAAGTGGCTGACCTTCTCT-----CTGGCAAAACAAAGTTGGGAACAAG 411
D 148 LysAsnTrpLeuHisPheGlnGlnLysCysTyrTyrPheGlyLys-----GlySerLys 165
QY 412 TTCCTCTGACCAATGCTGAAATATGACCTTTGAAAAAGTGAAGGCTTGTGTGTCAG 471
D 166 GlnTrpIle-----GlnAlaArgPheAlaCysSerAsp 176
QY 472 TTCAGGCTCTGTGGCCACCCAGGAATGTCAGAGATGGA---GCCATTGAGAAT 528
D 177 LeuGlnGlyArgLeuValSerIleHisSerGlnLysGluGlnAspPheLeuMetGlnHis 196
QY 529 CTCATCAAGGAGGAGGCTTCTCTGGGATCCTGAGAGAGAGAGAGAGAGGCGAGTTGTG 588
D 197 IleAsnLysLysAspSerTrpIleGlyLeuGlnAspLeuAsnMetGluGlyGluPheVal 216
QY 589 GATCTGACAGAAATAGACTGACCTACACAACTGGAACGAGGCTGAACCAACAATGCT 648
D 217 TrpSerAspGlySerProValGlyTyrSerAsnTrpAsnProGlyGluProAsnGly 236
QY 649 GGTTCCTGATGAAGTGTGTTGCTACTGAAATGCGCCAGTGAATGAGTCCCTGTCG 708
D 237 GlyGlnGlyGluAspCysValMetMetArgLysSerGlyGlnTrpAsnAspAlaPheCys 256
QY 709 TCCACC-----TCCCATCTGGCGCTCTGTGAG 735
D 257 ArgSerTyrLeuAspAlaTrpValCysGluGlnLeuAlaThrCysGlu 272

RESULT 10
US-08-365-103B-6
; Sequence 6, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G.
; APPLICANT: Nunez, Raphael D.
; APPLICANT: Yodoi, Jungi
; TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
; TITLE OF INVENTION: and Methods of Use for Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Ave. Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/365,103B
; FILING DATE: 28-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: Uirf N5-24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 288-3667
; TELEFAX: (515) 288-1338
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-365-103B-6

Alignment Scores:
Pred. No.: 4,12e-12 Length: 300
Score: 201.50 Matches: 51
Percent Similarity: 46.59% Conservative: 31
Best Local Similarity: 28.98% Mismatches: 47
Query Match: 14.76% Indels: 47
DB: 1 Gaps: 7

US-10-054-536-2 (1-747) x US-08-365-103B-6 (1-300)
QY 316 AGCTGCTGCTCCTCAGAAAGAAAGCT-----CTGCAACAGAA 354
D 121 SerLeuGlyLeuAsnGluLysArgThrAlaSerAspSerLeuGluLysLeuGlnGlu 140
QY 355 ATGCACGATC----- 366
D 141 ValAlaLysLeuTrpIleGluLeuSerLysGlyThrAlaCysAsnIleCysPro 160
QY 367 AAAAAGTGGCTGACCTTCTCT-----CTGGCAAAACAAAGTTGGGAACAAG 411
D 161 LysAsnTrpLeuHisPheGlnGlnLysCysTyrTyrPheGlyLys-----GlySerLys 178
QY 412 TTCCTCTGACCAATGCTGAAATATGACCTTTGAAAAAGTGAAGGCTTGTGTGTCAG 471
D 179 GlnTrpIle-----GlnAlaArgPheAlaCysSerAsp 189
QY 472 TTCAGGCTCTGTGGCCACCCAGGAATGCTGAGAGATGGA---GCCATTGAGAAT 528
D 190 LeuGlnGlyArgLeuValSerIleHisSerGlnLysGluGlnAspPheLeuMetGlnHis 209
QY 529 CTCATCAAGGAGGAGGCTTCTCTGGGATCCTGAGAGAGAGAGAGAGGCGAGTTGTG 588
D 210 IleAsnLysLysAspSerTrpIleGlyLeuGlnAspLeuAsnMetGluGlyGluPheVal 229
QY 589 GATCTGACAGAAATAGACTGACCTACACAACTGGAACGAGGCTGAACCAACAATGCT 648
D 230 TrpSerAspGlySerProValGlyTyrSerAsnTrpAsnProGlyGluProAsnGly 249
QY 649 GGTTCCTGATGAAGTGTGTTGCTACTGAAATGCGCCAGTGAATGAGTCCCTGTCG 708
D 250 GlyGlnGlyGluAspCysValMetMetArgLysSerGlyGlnTrpAsnAspAlaPheCys 269
QY 709 TCCACC-----TCCCATCTGGCGCTCTGTGAG 735
D 270 ArgSerTyrLeuAspAlaTrpValCysGluGlnLeuAlaThrCysGlu 285

RESULT 11
US-08-365-103B-2
; Sequence 2, Application US/08365103B
; Patent No. 5766943
; GENERAL INFORMATION:
; APPLICANT: Lynch, Richard G.
; APPLICANT: Nunez, Raphael D.
; APPLICANT: Yodoi, Jungi
; TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
```


Qy	487	GCACCCCCAGGAATGCTGCAGAGAGAAATGAGCCATTCAGAAATCTCATC----	AAGGAGGAA	5433
Db	203	AlaSerIleHisSerGlnGluGlnAspPheLeuAlaArgTyrAlaAsnLysIysGly		222
Qy	544	GCCTTCCTGGGCATCCTCATGATCAGAGACAGAGAGGCGAGTTGTGGATCTGCACGAGAAT		603
Db	223	ThrTrpIleGlyLeuArgAspLeuAspArgGluGlyGluPheIleTrpMetAspGluAsn		242
Qy	604	AGACTGACTACACAAACTGGAAACGAGGGTGAACCCCAACTGCTGGTTCATGAAGAT		663
Db	243	ProLeuAsnTyrSerAsnTrpArgProGlyGluProAsnAsnGlyGlyGlnGlyGluAsp		262
Qy	664	TGTTATTGCTACTGAAATAATGGCCAGTGGAAATGACGTCCCTCGTCCACCTCC-----		717
Db	263	CysValMetMetGlnGlySerGlyGlnTrpAsnAspAlaPheCysGlySerSerLeuAsp		282

QY	664	TGTTATATGCTAC	GGCAAAATGGCC	CAGGGAATGAC	GTCCCTGCTCC	ACCTCC-----	718
Db	263	CysValMetGln	GlySerGly	GlnTrpAsn	AspAlaPhe	CysGlySerSerLeu	Asp 282
QY	718	-----	CATCTGGCCGT	CTGT	732		
Db	283	GlyTrpValCys	AspArgLeu	AlaThrCys	292		

```

; Sequence 5, Application US/09535521
; Patent No. 6410714
;
; GENERAL INFORMATION:
;
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: AU-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24

```

```

; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 292
; TYPE: prt
; ORGANISM: Canis familiaris
US-09-535-521-5

```

Alignment Scores:	
Pred No :	1 220-11 000
Length:	000

Score:	136.50	Matches:	67
Percent Similarity:	44.35%	Conservative:	35
Best Local Similarity:	29.13%	Mismatches:	77
Query Match:	14.40%	Indels:	52

```
DB: 4
Caps: 11
```

US-10-054-536-2 (1-747) x US-09-535-521-5 (1-292)

QY 122 CAGGCATCAACGGCTTCCACAGCAAGATGGCGGTGATGCCACCAAGGAGAAAGGGG 181

Db 88 GlnAlaGlnValSerGlnAspMetLysGluIleGlnAla-----GluGlnLysArg 105

QY 182 AACGAGCCAGGGCTCAGAG-----GTTACAGGGCCCCC----- 217

Db 106 MetLysAlaGlnAspSerGluLeuSerGlnAsnLeuAspAlaLeuArgSerAspLeuAsn 125

QY 218 ---CTGGAAAGTTGGGGCCCTCCAGGAAATCCAGGGCCTTCTGGGTCAACGACCAAGG 274

Db 126 AsnLeuLysSerGlnSerLeuAsnGluArgSerThrAlaLeu--HisSerLeuGluArg 144

QY 275 GCCAAAAGGAGACCCCTGGAAAAGTCCGGATGGTGATAGTACCTGGCTGCCCTCAGAAA 334

Db
145 LeuGlnGluGlu---ValGluLysLeu-----TrpMet----- 154

QY 335 GAAAGCTCTGCAACAGAAATGGCACGTATCAA----- 369

Db 155GluLeuHisValSerAsnGlySerGluCysAsnThrCysPro 168

QY 370 ---AAGTGGCTGACCTTCTCTCTGGGCAACAAGTTGGGAACAAGTTCTTCTGACCAAT 426

Blank